

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 14:10:20 ; Search time 2707 Seconds
(without alignments)
7994.530 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529
Sequence: 1 catgtgtatgtgttagtcgt.....tcttggagagggcttcttgag 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.ey.*

13: gb.un.*

14: gb.vi.*

15: gb.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rodi.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.ey.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	517.6	97.8	117978	9	AC079200	AC079200 Homo sapi
2	517.6	97.8	167118	9	AC023385	AC023385 Homo sapi
3	451.6	85.4	63888	2	AC084244	AC084244 Homo sapi
4	58.4	11.0	341050	3	PFA929357	AL929357 Plasmid
5	52.4	9.9	251551	3	AE014844	AE014844 Plasmid
6	50.8	9.6	146570	3	AC117072	AC117072 Dictyoste
7	50.6	9.6	958	3	DDI1092A	M19468 D.discoideu
8	50.6	9.6	2255	3	DDITUBA	L13999 Dictyosteli
9	50.6	9.6	331039	3	AC116988	AC116988 Dictyoste
10	50.4	9.5	110000	2	PFMAL8P1_01	Continuation (2 of
11	49.8	9.4	16744	2	EX469924	EX469924 Danio rer
12	49.6	9.4	518	3	AY221171	AY221171 Dictyoste
13	49.4	9.3	256879	3	AC116982	AC116982 Dictyoste
14	49.2	9.3	5850	3	DDIDDP2	M55298 Dictyosteli
15	49.2	9.3	5852	3	DDP2PLAS	X51478 Dictyosteli
16	48.2	9.3	302156	3	AZ1622	AZ1622 Dictyosteli
17	48.8	9.2	302156	3	AC116977	AC116977 Dictyoste
18	48.6	9.2	611	6	AX187051	AX187051 Sequence
19	48.6	9.2	343050	3	PFA929353	AL929353 Plasmid
20	48.4	9.1	1052	3	AF020696	AF020696 Trypanoso
21	48.4	9.1	253305	3	PFMAL3P7	AL034559 Plasmid
22	48.2	9.1	39989	2	AC091136	AC091136 Homo sapi
23	48.2	9.1	146150	2	AC021298	AC021298 Homo sapi
24	48.2	9.1	155985	9	AC124657	AC124657 Homo sapi
25	48.2	9.1	181996	2	AC068749	AC068749 Homo sapi
26	48.2	9.1	256221	2	AL356371	AL356371 Homo sapi
27	47.8	9.0	36188	3	AC115608	AC115608 Dictyoste
28	47.8	9.0	54441	3	AC115584	AC115584 Dictyoste
29	47.8	9.0	152336	2	AC009594	AC009594 Homo sapi
30	47.8	9.0	162554	9	AC079354	AC079354 Homo sapi
31	47.8	9.0	182012	9	AL359846	AL359846 Human DNA
32	47.6	9.0	6457	3	AC115613	AC115613 Dictyoste
33	47.6	9.0	61052	2	AC123513	AC123513 Dictyoste
34	47.6	9.0	136240	3	AC117070	AC117070 Dictyoste
35	47.6	9.0	163443	2	AC006280	AC006280 Plasmid
36	47.6	9.0	196149	2	AC004709	AC004709 Plasmid
37	47.6	9.0	250029	3	AE014830	AE014830 Plasmid
38	47.6	9.0	252650	3	AE014847	AE014847 Plasmid
39	47.4	9.0	110000	3	PFMAL1P2_1	Continuation (2 of
40	47.4	9.0	113515	6	AX347077	AX347077 Sequence
41	47.4	9.0	156060	2	AC004153	AC004153 Plasmid
42	47.4	9.0	176109	9	AC003992	AC003992 Homo sapi
43	47.4	9.0	176379	9	AC016751	AC016751 Homo sapi
44	47.4	9.0	250707	3	AE014848	AE014848 Plasmid
45	47.4	9.0	253305	3	PFMAL3P7	AL034559 Plasmid
46	47.2	8.9	2992	3	AC112945	AC112945 Dictyoste
47	47.2	8.9	111882	3	AC115612	AC115612 Dictyoste
48	47.2	8.9	340552	3	PFA929354	AL929354 Plasmid
49	47.2	8.9	611	6	AX187051	AX187051 Sequence
50	47.2	8.9	24479	3	AC114265	AC114265 Dictyoste
51	47.2	8.9	105682	3	AC116957_3	Continuation (4 of
52	46.8	8.8	2245	3	AV160090	AV160090 Dictyoste
53	46.8	8.8	12029	3	AE001427	AE001427 Plasmid
54	46.8	8.8	110000	2	PFMAL7P1_02	Continuation (3 of
55	46.8	8.8	118587	2	AC140024	AC140024 Medicago
56	46.8	8.8	140884	9	AC139618	AC139618 Homo sapi
57	46.8	8.8	252420	3	AE014841	AE014841 Plasmid
58	46.8	8.8	273275	3	AE014828	AE014828 Plasmid
59	46.8	8.8	343050	3	PFA929353	AL929353 Plasmid
60	46.6	8.8	82078	9	AC020379	AC020379 Homo sapi
61	46.6	8.8	110000	2	AL732359_09	Continuation
62	46.6	8.8	110000	3	AC116984_3	Continuation
63	46.6	8.8	166420	2	AC012112	AC012112
64	46.6	8.8	252394	3	AE014833	AE014833
65	46.4	8.8	56080	8	AP003838	AP003838

66	46.4	8.8	93459	8	AP006086	AP006086 Lotus jap
67	46.4	8.8	107764	8	AP004984	AP004984 Lotus jap
68	46.4	8.8	110000	2	PFMAL13_26	Continuation (27 of
69	46.4	8.8	110000	2	PFMAL7P1_06	Continuation (7 of
70	46.4	8.8	110000	3	AC116957_1	Continuation (2 of
71	46.4	8.8	192187	3	AC116920_1	AC116920 Dictyoste
72	46.4	8.8	192187	3	AC116920_1	AC116920 Dictyoste
73	46.2	8.7	12419	3	AE001379	AE001379 Plasmodiu
74	46.2	8.7	162776	2	BX511089	BX511089 Danio rer
75	46.2	8.7	164399	3	PFMAL3P6	Z98551 Plasmodium
76	46.2	8.7	172559	2	AC101795	AC101795 Mus muscu
77	46.2	8.7	263434	2	AC135774	AC135774 Rattus no
78	46.2	8.7	273275	3	AE014828	AE014828 Plasmodiu
79	46.2	8.7	332526	2	BX255919	BX255919 Danio rer
80	46	8.7	6281	2	AF465307	AF465307 Dictyoste
81	46	8.7	26203	2	PFMAL3P7	AF096784 Plasmodiu
82	46	8.7	48063	8	AF288090	AF288090 Rhodomon
83	46	8.7	68730	2	AC100664	AC100664 Mus muscu
84	46	8.7	114665	2	PFMAL7P1_02	Continuation (3 of
85	46	8.7	146570	3	AC113089	AC113089 Medicago
86	46	8.7	151802	3	AC111702	AC111702 Dictyoste
87	46	8.7	164135	2	AC114263	AC114263 Dictyoste
88	46	8.7	232323	2	AC102228	AC102228 Mus muscu
89	46	8.7	251448	3	AC127590	AC127590 Mus muscu
90	46	8.7	266544	3	AE014819	AE014819 Plasmodiu
91	46	8.7	349980	6	AC116956	AC116956 Dictyoste
92	45.8	8.7	660	3	AX344563	AX344563 Sequence
93	45.8	8.7	40611	3	AF440515	AF440515 Caenorhab
94	45.8	8.7	110000	2	AC116987	AC116987 Dictyoste
95	45.8	8.7	110000	2	PFMAL7P1_05	Continuation (6 of
96	45.8	8.7	13880	3	PFMAL3P4	AL008970 Plasmodiu
97	45.8	8.7	126876	9	HSDJ34M23	AL121988 Human DNA
98	45.8	8.7	170027	2	AC110670	AC110670 Canis fam
99	45.8	8.7	187685	2	BX005107	BX005107 Danio rer
100	45.8	8.7	241614	2	AC114121	AC114121 Rattus no
					AE014834	AE014834 Plasmodiu

RESULT 1	AC079200	11978 bp	DNA	linear	PRI 01-MAY-2002
LOCUS	AC079200	11978 bp	DNA	linear	PRI 01-MAY-2002
DEFINITION	AC079200	11978 bp	DNA	linear	PRI 01-MAY-2002
ACCESSION	AC079200	11978 bp	DNA	linear	PRI 01-MAY-2002
VERSION	AC079200.6	GI:20377011			
KEYWORDS	HTG.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Birren,B., Linton,L., Nusbaum,C., Lander,E., Brown,A.,				
JOURNAL	Choeplai,Y., Colangelo,M., Collins,S., Collymore,A.,				
REFERENCE	Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,				
AUTHORS	FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S.,				
	Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,				
	Iliev,I., Johnson,R., Jones,C., Kam,L., Karatas,A., LaRocque,K.,				
	Lamazares,R., Lander,E., Lehoucq,J., Levine,R., Lieu,C., Liu,G.,				
	Macdonald,P., Marquis,N., McEwan,P., Meneus,L., Mihova,T.,				
	Morow,J., Murphy,T., Nguyen,J., Norman,C.H., O'Connor,T.,				
	O'Donnell,P., O'Neill,D., Pollara,V., Riese,C., Rogov,P.,				
	Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,				
	Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,				
	Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,				
	Viell,J., Vo,A., Willson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,				
	Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.				

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-105B24"
/clone_lib="RPC1-11 Human Male BAC"
complement(2..45)
/rpt_family="THE1B"
413..435
/rpt_family="AT_rich"
442..4918
/rpt_family="LIMAS"
1031..1232
/rpt_family="LIMB3"
complement(1386..1622)
/rpt_family="MIR"
1782..2075
/rpt_family="AluSq"
2157..2407
/rpt_family="LIPB1"
2408..2437
/rpt_family="(TG)n"
2437..2461
/rpt_family="(TA)n"
2462..2827
/rpt_family="LIPB1"
3113..3144
/rpt_family="(TG)n"
complement(5098..5729)
/rpt_family="LIMB8"
6062..6209
/rpt_family="LIM4"
6591..6668
/rpt_family="MER116"
complement(7056..7171)
/rpt_family="MIR"
7595..7933
/rpt_family="MLT1J2"
complement(8371..8498)
/rpt_family="LIMAS"
complement(8782..8958)
/rpt_family="MIR3"
10392..10495
/rpt_family="MIR"
complement(10523..10550)
/rpt_family="MIR"
10620..10688
/rpt_family="MIR"
11359..11423
/rpt_family="(TGGA)n"
11445..11611
/rpt_family="(TGGA)n"
complement(11724..11993)
/rpt_family="L3"
12093..12262
/rpt_family="LTR41"
complement(12751..12896)
/rpt_family="MIR"
12903..13006
/rpt_family="HAL1"
complement(13397..13512)
/rpt_family="L2"
13613..14101
/rpt_family="MLT2B1"
14267..14555
/rpt_family="AluSp"
complement(14845..15049)
/rpt_family="LIMC5"
complement(15054..15164)
/rpt_family="LIMC5"
16324..16627
/rpt_family="AluY"

repeat_region complement(17668..17733)
/rpt_family="L2"
repeat_region complement(18517..18621)
/rpt_family="MIR"
repeat_region 19516..20673
/rpt_family="LIME1"
repeat_region 20701..20841
/rpt_family="MIR3"
repeat_region 23592..24723
/rpt_family="LIMC2"
repeat_region complement(24737..24859)
/rpt_family="FLAM_C"
repeat_region 24904..24967
/rpt_family="L2"
repeat_region complement(27155..27449)
/rpt_family="AluSx"
repeat_region complement(27756..27864)
/rpt_family="L2"
repeat_region 28024..28283
/rpt_family="MIR"
repeat_region 28624..28864
/rpt_family="AluSq"
repeat_region 29251..29273

Query Match 97.8%; Score 517.6; DB 9; Length 117978;
Best Local Similarity 99.6%; Pred. No. 3e-104;
Matches 528; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CATGGTGATGCTGTAGTCGGTGGCTGATTTATATGCTGATTTATGCTGATTTATGGTGATTTGCTTCT 60
DB 88017 CATGGTGATGCTGTAGTCGGTGGCTGATTTATATGCTGATTTATGGTGATTTGCTTCT 88076
QY 61 TCTTTATACCTTTTATTTATTTCCCAAAATTTTCTTAAAGCAAAATATTTCTTGGCTAATCAAT 120
DB 88077 TCTTTATACCTTTTATTTATTTATTTCCCAAAATTTTCTTAAAGCAAAATATTTCTTGGCTAATCAAT 88136
QY 121 AAATTTATCAAAAGAAAAAAGAACTGAAAGCAACGCTTGAAAGAAAGGAAAGTTAG-CCCTA 179
DB 88137 AAATTTATCAAAAGAAAAAAGAACTGAAAGCAACGCTTGAAAGAAAGGAAAGTTAGCCCTA 88196
QY 180 TCGGGTATATTTTGGAAAGTTGTAAATACTACTACTACTCTCTTCCATGACAAACCAACCATGAC 239
DB 88197 TCGGGTATATTTTGGAAAGTTGTAAATACTACTACTACTCTCTTCCATGACAAACCAACCATGAC 88256
QY 240 TTTTCTTTGACAGCAAGAAAGCAAGCAGTCACCTCTCTTCCATGACAAACCAACCATGAC 299
DB 88257 TTTTCTTTGACAGCAAGAAAGCAAGCAGTCACCTCTCTTCCATGACAAACCAACCATGAC 88316
QY 300 CGGCTGCTTGCTGCTCCATCACAGGAAGTGCAGAGAGCCCTTTTCTTCCAGGGTTG 359
DB 88317 CGGCTGCTTGCTGCTCCATCACAGGAAGTGCAGAGAGCCCTTTTCTTCCAGGGTTG 88376
QY 360 TGGGACCTTGCACTATCTGCTTGGCTTGTAGCTAGAACTGCGCCCAAGTGCAGCAGCTC 419
DB 88377 TGGGACCTTGCACTATCTGCTTGGCTTGTAGCTAGAACTGCGCCCAAGTGCAGCAGCTC 88436
QY 420 AAGCTGTTTCGGGGACTTCCAGATAAACCAGAACAGCTGTATTACTGTCTTTTGACATCC 479
DB 88437 AAGCTGTTTCGGGGACTTCCAGATAAACCAGAACAGCTGTATTACTGTCTTTTGACATCC 88496
QY 480 TCTCTGGTCAGGGCCAAAGTCACTGTCCTGCTTGGAGAGGGCTTCTGAG 529
DB 88497 TCTCTGGTCAGGGCCAAAGTCACTGTCCTGCTTGGAGAGGGCTTCTGAG 88546

RESULT 2
AC023385/c
LOCUS AC023385 167118 bp DNA linear PRI 31-JUL-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-675P19, complete sequence.
ACCESSION AC023385
VERSION AC023385.10 GI:22024594
KEYWORDS HTG.
SOURCE Homo sapiens (human)
```

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 (bases 1 to 167118)
Birren, B., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 8, clone RP11-675P19
Unpublished

TITLE
2 (bases 1 to 167118)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bada, F., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,
Fenster, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Landers, T., Largocque, K., Lehocsky, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Miranda, C., Menga, V., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A.,
Travers, M., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and
Zody, M.

TITLE
Direct Submission
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
3 (bases 1 to 167118)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Menga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE
Direct Submission
Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
4 (bases 1 to 167118)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Menga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1
* 656: contig of 656 bp in length
* 756: gap of 100 bp
* 757
* 1450: contig of 694 bp in length
* 1451
* 1550: gap of 100 bp
* 1551
* 2232: contig of 682 bp in length
* 2233
* 2332: gap of 100 bp
* 2333
* 2983: contig of 651 bp in length
* 2984
* 3083: gap of 100 bp
* 3084
* 3725: contig of 642 bp in length
* 3726
* 3825: gap of 100 bp
* 3826
* 4455: contig of 630 bp in length
* 4456
* 4555: gap of 100 bp
* 4556
* 5212: contig of 657 bp in length
* 5213
* 5312: gap of 100 bp
* 5313
* 5978: contig of 666 bp in length
* 5979
* 6078: gap of 100 bp
* 6079
* 6758: contig of 680 bp in length
* 6759
* 6858: gap of 100 bp
* 6859
* 7534: contig of 676 bp in length
* 7535
* 7634: gap of 100 bp
* 7635
* 8307: contig of 673 bp in length
* 8308
* 8407: gap of 100 bp
* 8408
* 9064: contig of 657 bp in length
* 9065
* 9164: gap of 100 bp
* 9165
* 9843: contig of 679 bp in length
* 9844
* 9943: gap of 100 bp
* 9944
* 10631: contig of 688 bp in length
* 10632
* 10731: gap of 100 bp
* 10732
* 11419: contig of 688 bp in length
* 11420
* 11519: gap of 100 bp
* 11520
* 12185: contig of 666 bp in length
* 12186
* 12285: gap of 100 bp
* 12286
* 12931: contig of 646 bp in length
* 12932
* 13031: gap of 100 bp
* 13032
* 13700: contig of 669 bp in length
* 13701
* 13800: gap of 100 bp
* 13801
* 14482: contig of 682 bp in length
* 14483
* 14582: gap of 100 bp
* 14583
* 15261: contig of 679 bp in length
* 15262
* 15361: gap of 100 bp
* 15362
* 16042: contig of 681 bp in length
* 16043
* 16142: gap of 100 bp
* 16143
* 16823: contig of 681 bp in length
* 16824
* 16923: gap of 100 bp
* 16924
* 17615: contig of 692 bp in length
* 17616
* 17715: gap of 100 bp
* 17716
* 18373: contig of 658 bp in length
* 18374
* 18473: gap of 100 bp
* 18474
* 19160: contig of 687 bp in length
* 19161
* 19260: gap of 100 bp
* 19261
* 19932: contig of 672 bp in length
* 19933
* 20032: gap of 100 bp
* 20033
* 20689: contig of 657 bp in length
* 20690
* 20789: gap of 100 bp
* 20790
* 21446: contig of 657 bp in length
* 21447
* 21546: gap of 100 bp
* 21547
* 22186: contig of 640 bp in length
* 22187
* 2286: gap of 100 bp
* 2287
* 22973: contig of 687 bp in length
* 22974
* 23073: gap of 100 bp
* 23074
* 23762: contig of 689 bp in length
* 23763
* 23862: gap of 100 bp
* 23863
* 24538: contig of 676 bp in length
* 24539
* 24638: gap of 100 bp
* 24639
* 25309: contig of 671 bp in length
* 25310
* 25409: gap of 100 bp

* 25410
* 26101: contig of 692 bp in length
* 26201: gap of 100 bp
* 26202
* 26893: contig of 692 bp in length
* 26894
* 26993: gap of 100 bp
* 26994
* 27685: contig of 692 bp in length
* 27686
* 27785: gap of 100 bp
* 27786
* 28456: contig of 671 bp in length
* 28457
* 28556: gap of 100 bp
* 28557
* 29193: contig of 637 bp in length
* 29194
* 29293: gap of 100 bp
* 29294
* 29970: contig of 677 bp in length
* 29971
* 30070: gap of 100 bp
* 30071
* 30757: contig of 687 bp in length
* 30758
* 30857: gap of 100 bp
* 30858
* 31520: contig of 663 bp in length
* 31521
* 31620: gap of 100 bp
* 31621
* 32313: contig of 693 bp in length
* 32314
* 32413: gap of 100 bp
* 32414
* 33074: contig of 661 bp in length
* 33075
* 33174: gap of 100 bp
* 33175
* 33880: contig of 706 bp in length
* 33881
* 33980: gap of 100 bp
* 33981
* 34667: contig of 687 bp in length
* 34668
* 34767: gap of 100 bp
* 34768
* 35390: contig of 623 bp in length
* 35391
* 35490: gap of 100 bp
* 35491
* 36141: contig of 651 bp in length
* 36142
* 36241: gap of 100 bp
* 36242
* 36873: contig of 632 bp in length
* 36874
* 36973: gap of 100 bp
* 36974
* 37642: contig of 669 bp in length
* 37643
* 37742: gap of 100 bp
* 37743
* 38426: contig of 684 bp in length
* 38427
* 38526: gap of 100 bp
* 38527
* 39190: contig of 664 bp in length
* 39191
* 39290: gap of 100 bp
* 39291
* 39592: contig of 662 bp in length
* 39593
* 40052: gap of 100 bp
* 40053
* 40712: contig of 660 bp in length
* 40713
* 40812: gap of 100 bp
* 40813
* 41470: contig of 658 bp in length
* 41471
* 41570: gap of 100 bp
* 41571
* 42240: contig of 670 bp in length
* 42241
* 42340: gap of 100 bp
* 42341
* 43014: contig of 674 bp in length
* 43015
* 43114: gap of 100 bp
* 43115
* 43762: contig of 648 bp in length
* 43763
* 43862: gap of 100 bp
* 43863
* 44496: contig of 634 bp in length
* 44497
* 44596: gap of 100 bp
* 44597
* 45274: contig of 678 bp in length
* 45275
* 45374: gap of 100 bp
* 45375
* 46071: contig of 697 bp in length
* 46072
* 46171: gap of 100 bp
* 46172
* 46850: contig of 679 bp in length
* 46851
* 46950: gap of 100 bp
* 46951
* 47625: contig of 675 bp in length
* 47626
* 47725: gap of 100 bp
* 47726
* 48419: contig of 694 bp in length
* 48420
* 48519: gap of 100 bp
* 48520
* 49175: contig of 656 bp in length
* 49176
* 49275: gap of 100 bp
* 49276
* 49967: contig of 692 bp in length
* 49968
* 50067: gap of 100 bp
* 50068
* 50752: contig of 685 bp in length
* 50753
* 50852: gap of 100 bp
* 50853
* 51504: contig of 652 bp in length
* 51505
* 51604: gap of 100 bp
* 51605
* 52254: contig of 650 bp in length
* 52255
* 52354: gap of 100 bp
* 52355
* 53006: contig of 652 bp in length
* 53007
* 53106: gap of 100 bp
* 53107
* 53786: contig of 680 bp in length

Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
Direct Submission
JOURNAL Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES
Location/Qualifiers

source 1..251551
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"

repeat_region 4..647
/rpt_type=tandem
repeat_region 656..797
/rpt_type=tandem
repeat_region 754..933
/rpt_type=tandem
repeat_region 775..807
/rpt_type=tandem
repeat_region 2009..2796
/rpt_type=tandem
repeat_region 2868..3434
/rpt_type=tandem
repeat_region 3622..3748
/rpt_type=tandem
repeat_region 8569..8734
/rpt_type=tandem
repeat_region 9066..9276
/rpt_type=tandem
repeat_region 9086..9140
/rpt_type=tandem
repeat_region 9220..9262
/rpt_type=tandem
repeat_region 9374..9469
/rpt_type=tandem
repeat_region 9374..9452
/rpt_type=tandem
repeat_region 9550..9629
/rpt_type=tandem
repeat_region 9550..9601
/rpt_type=tandem
repeat_region 9550..9593
/rpt_type=tandem
repeat_region 9670..10074
/rpt_type=tandem
repeat_region 9694..9746
/rpt_type=tandem
repeat_region 9752..10260
/rpt_type=tandem
repeat_region 9837..9899
/rpt_type=tandem
repeat_region 9891..10248
/rpt_type=tandem
repeat_region 10189..10254
/rpt_type=tandem
repeat_region 10243..11081
/rpt_type=tandem
repeat_region 10893..10944
/rpt_type=tandem
repeat_region 11078..12495
/rpt_type=tandem
repeat_region 11078..12488
/rpt_type=tandem
repeat_region 11078..11773
/rpt_type=tandem
repeat_region 11246..11306
/rpt_type=tandem
repeat_region 11916..11963
/rpt_type=tandem
repeat_region 12061..12685
/rpt_type=tandem
repeat_region 12744..13089
/rpt_type=tandem

repeat_region 12744..13048
/rpt_type=tandem
repeat_region 12828..13048
/rpt_type=tandem
repeat_region 12995..13130
/rpt_type=tandem
repeat_region 13041..13636
/rpt_type=tandem
repeat_region 13049..13253
/rpt_type=tandem
repeat_region 13073..13298
/rpt_type=tandem
repeat_region 13269..13560
/rpt_type=tandem
repeat_region 13584..13760
/rpt_type=tandem
repeat_region 13617..13819
/rpt_type=tandem
repeat_region 13662..14053
/rpt_type=tandem
repeat_region 13954..14071
/rpt_type=tandem
repeat_region 13961..14088
/rpt_type=tandem
repeat_region 14191..14303
/rpt_type=tandem
repeat_region 14309..14357
/rpt_type=tandem
repeat_region 14467..14669
/rpt_type=tandem
repeat_region 14474..14581
/rpt_type=tandem
repeat_region 14571..14689
/rpt_type=tandem
repeat_region 14590..14686
/rpt_type=tandem
repeat_region 14744..14880
/rpt_type=tandem
repeat_region 14917..14944
/rpt_type=tandem
repeat_region 15059..15091
/rpt_type=tandem
repeat_region 15143..15246
/rpt_type=tandem
repeat_region 15862..16005
/rpt_type=tandem
repeat_region 16738..16826
/rpt_type=tandem

gene

mRNA join(<16973..22048,23037..>24497)

CDS locus_tag="PFL0005w"

note="HMPFam hit to PF03011, Plasmodium falciparum

erythrocyte membrane protein (pFEMP), score 1.4e-64"

/codon_start=1

/product="erythrocyte membrane protein 1 (pFEMP1)"

/protein_id="AAN36090.1"

/db_xref="GI:23496532"

/translation="MGQPAVTDYSKATNVKYLFDLIGETVQKAKAKADVASKRYFE
ELHGLSKATYKDDKEGTTTPNPCKLEYQYHTNVTGKEDKYPCLGRKTVRFSDKE
GACVYKTKIDSTTDTVGACAPYRRLVYMCNRLSHIPTKITTNLLLDVCLAQYEG
QSIQNHGHQLSYSPDSOLCTELARFADIGDIVRGDLVYRGNRNDKLEKLLKG
YFKKIDYLNVEKKKEAEATYKDDADPFYQLRDEWALNRQDVWKAITCDADHSYRK
MGADGSIYESAMRQCRNADVPTNFDPVQYLRWFBEWAEDFCRRKRGKJLKDATQKCR
QDGTGRDYCDLNRDKRTISARHELQVBECKGCVICIPFGWIDNQBFKQ
KNKYTNINKKHDETTKEISGNRRKRSLTINKYGYDEEFYKIPKBYIPDVQKFLDL
LSKETACESQYDEPRTTISINFKYKNPDI FSHTEYCOACPMCGMTCTFDGCKTKNPE
ELCHHKIVQKEYPDTNTDIPILTPDTTKSNIVKRYNPNCSDDNNSDQINNQOCHY
DESKSGNONCVETQWQNFQKQKVTSYNAFFPWKVMSELDSDSIKWRBGLDKLND
KKTGCKKCRNDCKCFKKWVEQKKEKWKAIKGHPKKQKQMIETGMPPEMAKLKLLND
VFLQDMBEKAQGDPOHIAIKIKELLKKNDKVNNLNMMETIFDILLQEBEQDAQKCVSNV

/protein_id="AAA02987.1"
/db_xref="GI:167574"
/translations="MNIRSLIILIIISTILFFSIINGSLSLDPTCVGAPDQGVYLRSSM
DFKGRVYVNVSGYLSLSDGFRGNVQSFISGADVCFVXWPIEQYQITAGESHRNYA
ALTNFGQMDAILIPGNSNIVCSFK"
BASE COUNT 380 a 90 c 95 g 393 t
ORIGIN Unreported.

Query Match 9.6%; Score 50.6; DB 3; Length 958;
Best Local Similarity 53.3%; Pred. No. 0.67; Indels 0; Gaps 0;
Matches 104; Conservative 1; Mismatches 90; Indels 0; Gaps 0;
QY 47 TGAATTTGCTCTCTTATACCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATATT 106
Db 104 TTAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 163
QY 107 CTTTGCTAATCAATTAATTAATCAAGAAAAAAGAACTGAAGCAAGCTTGAAAAAGG 166
Db 164 TTTTATTTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAG 223
QY 167 AAGATTAGCCCTATCGGTATATTTTGGAGTTTGTAAATTAATTAATTAATTAATTAATTAAG 226
Db 224 AAGTCGGTTAAATAAATCGAATTTTATTTTAAAGTTTATTTTAAATTAATTAATTAAG 293
QY 227 TCCCACTCCTCTGTT 241
Db 284 TAATATTGATTTT 298

RESULT 8
DDITUBA 2255 bp DNA linear INV 21-SEP-1993
LOCUS Dictyostelium discoideum alpha tubulin (TUBA) gene, complete cds.
DEFINITION L13939
ACCESSION L13939
VERSION L13939.1 GI:290058
KEYWORDS alpha-tubulin.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 2255)
AUTHORS Trivinos-Lagos, L., Omachi, T., Albrightson, C., Burns, R.G.,
Ennis, H.L., and Chisholm, R.L.
TITLE The highly divergent alpha- and beta-tubulins from Dictyostelium
discoideum are encoded by single genes
J. Cell. Sci. 105 (Pt 4), 903-911 (1993)
JOURNAL 94043557
MEDLINE 8227212
PUBMED
COMMENT Original source text: Dictyostelium discoideum (strain AX-3)
(library: Clontech) DNA.
FEATURES
source
1. .2255
/organism="Dictyostelium discoideum"
/strain="AX-3"
/db_xref="taxon:44689"
/tissue_lib="Clontech"
178. .1662
/gene="TUBA"
join(178. .805, 917. .1662)
/gene="TUBA"
/note="putative"
/codon_start=1
/product="alpha tubulin"
/protein_id="AAC37343.1"
/db_xref="GI:290059"

/translation="MGREIIISHIGQVGVGNSCWELLYCLBHGIERDGSIPADRKQS
SDVNNKLGITFFSESTGKVVPRAFIDLEPTVIDEIRTDYKNLPHPLQITGK
EAAANVARGHTVGVKELIDVCDVIRRLADQDGLQGLVFLHVSFGTSGFGLLL
OKALDYGGKSKLDFCVPSQVSVSVVPEVNSVLSLSHLESDVSMLEALYN
ICKNSVLEKPTVNLRLIAOVLSLTSRLFPGLNDINDIOTNLVPFRLHFLV
CSVAPVIREKAHETITVDNITSVFESEKNLAKOPNLGKYMACKLWYRGDIVPKE
AQKAVQNIRESKRVNSVDMSPTEGKCGINNAQPVSTKDSMAEVKSKVCMLSNTTA
ISGVFSRINHKKDLMEVKKRAVHWVGEWEGEFAERDILLALEKDYESVSASTEG
EQEEBEY"

exon <178. .805
/gene="TUBA"
intron 806. .916
/gene="TUBA"
exon 917. .>1662
/gene="TUBA"
BASE COUNT 823 a 324 c 283 g 825 t
ORIGIN
Query Match 9.6%; Score 50.6; DB 3; Length 2255;
Best Local Similarity 53.3%; Pred. No. 0.65; Indels 0; Gaps 0;
Matches 104; Conservative 1; Mismatches 90; Indels 0; Gaps 0;
QY 47 TGAATTTGCTCTCTTATACCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATATT 106
Db 2061 TTAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2120
QY 107 CTTTGCTAATCAATTAATTAATCAAGAAAAAAGAACTGAAGCAAGCTTGAAAAAGG 166
Db 2121 TTTTATTTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAG 2180
QY 167 AAGATTAGCCCTATCGGTATATTTTGGAGTTTGTAAATTAATTAATTAATTAATTAAG 226
Db 2181 AAGTCGGTTAAATAAATCGAATTTTATTTTAAAGTTTATTTTAAATTAATTAATTAAG 2240
QY 227 TCCCACTCCTCTGTT 241
Db 2241 TAATATTGATTTT 2255

RESULT 9
AC116988/c 331039 bp DNA linear INV 12-MAR-2003
LOCUS Dictyostelium discoideum chromosome 2 map 6445720-6776760 strain
DEFINITION AX4, complete sequence.
ACCESSION AC116988
VERSION AC116988.2 GI:28829381
KEYWORDS HTG.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 331039)
AUTHORS Gloeckner, G., Eichinger, L., Szafranski, K., Pachebat, J., Dear, P.,
Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K.,
Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and
Noegel, A.A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
Nature 418 (6893), 79-85 (2002)

JOURNAL MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
REFERENCE 2 (bases 1 to 331039)
AUTHORS Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 331039)
AUTHORS Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 331039)
AUTHORS Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On Mar 4, 2003 this sequence version replaced gi:20042960.
CDS predictions from GeneID do not necessarily reflect true genes.
Further Information is available from IMB Jena, Department of
Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml)


```

QY      3  TGGTGAAGCTAGTGCCTGGCTGATTTATATGCTGATTTATGCGGATTTTCTTCCTTC 62
Db      169225 TGTATTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 169284

QY      63  TTTATATCTTTTATTTATCCCAATTTTCTTAAGCAATATTTCTTTGCTATCAATAA 122
Db      169285 TTTATTTATTTATTTATTTATTTTATTTTATTTTATTTTAAAGTCACAGAGAGAAAGAAA 169344

QY      123 ATTATCAAAAGAAAAGAAAAGCTGAAACCAACGCTTGAAAAAGGAAAGTATGCCCTATCG 182
Db      169345 AAAAAAGAAAAGAAAAGAAAAGTAAAAAGGAAAGAAAAGAAAAGAAAAGAAAAGTGA 169404

QY      183 GGTATATTTTGAAGTTGTAAAA 205
Db      169405 AAAAGATTTTAAATATATCCAAAA 169427

RESULT 14
DDIDDP2/c
LOCUS
DEFINITION
Dictyostelium discoideum plasmid Ddp2 trans-acting factor gene,
complete cds.
ACCESSION
M55298
VERSION
M55298.1 GI:167727
KEYWORDS
trans-acting factor.
SOURCE
Dictyostelium discoideum
ORGANISM
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE
1 (bases 1 to 5850)
AUTHORS
Leiting,B., Lindner,I.J. and Noegel,A.A.
TITLE
The extrachromosomal replication of Dictyostelium plasmid Ddp2
requires a cis-acting element and a plasmid-encoded trans-acting
factor
JOURNAL
Mol. Cell. Biol. 10 (7), 3727-3736 (1990)
MEDLINE
90287164
PUBMED
2192261
COMMENT
Original source text: Dictyostelium discoideum (strain WS380B) DNA.
FEATURES
     source
       1..5850
         /organism="Dictyostelium discoideum"
         /mol_type="genomic DNA"
         /strain="WS380B"
         /db_xref="taxon:44689"
         /note="putative VECTOR sequence Bacteriophage lambda
(J02459); putative"
         1795..1829
           /codon_start=1
           /product="trans-acting factor"
           /protein_id="AAA33191.1"
           /db_xref="GI:167728"
           /translation="MDELISMDRFFKFFVILLKEFKGKRNDRVLSVDYDILGIYSP
RTFVLKEVPRAVAVSYDESEIDLFRIGVPEPCTSLYSYIPGIFSLKDFLLISKTSKG
IRVSDVDQAILLDFHRSRISDKQVRKDIIPGRTRPEKLSSEYKISDGRAGVSWFN
LVKSLSTCKNHLFAENTYKHFDFISMLSLVHGIIVDSQNEENNVSAMVSLNPFV
DLEKSDIPGAVSRVTNTRGSRNSLNLPFTTTTTTTTTTAPITTRSKRSDD
EIGLEWVKVLQDQKYSTGLLENFNHNSERVITLSSSFFEIASKINDYEVKFS
ELSIDVLESAKLTREKNTNLIPTNNKEGEPFLWPLVNGIATSVFVSNNYSVG
SPANVESAKLHLICISGNINGFSLRSITPDTFKSITKDLIPMSKRMLDLEQGRK
LRDANNNGKSKQSDISGIDETDKLISFVHEFINDLNLYLKSKEEDGLMLVDFP
TSLFMRYNPNISDNKGVMPHCRSEISFQCKNHSIDNLVLSFTPNKINISQDNE
NELKKYSLMVSDFRNPKVTPKPISEPFRFTIITFTNNSYNANRVAFDDISGIS
ITNKNIIHAKGORNFEIYTLAGSTRIRAFPCALQIINNFKATDKLIDDOSVNH
QIASLEIKNLSVLPDLKVRGTVGKGETAPIIINSEDTFTSICLDIRFSASLI
SKTKLSQLTFPAPRYNKETNLIKVDQCBELTFTFLNNYKIANKLSTIENLYNNF
MGLEDEDEDEDEDEDEDEDEDEDEG"
       complement(5073..5113)
       /note="putative VECTOR sequence Cloning vector pEMBL 8
minus (pEMBL8m) (X04996); putative"
       2300 a 646 c 706 g 2198 t
BASE COUNT 2300 a 646 c 706 g 2198 t
ORIGIN
Plasmid Ddp2.

```

```

Query Match          9.3%; Score 49.2; DB 3; Length 5850;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY      12  TGTAGTCGTGCTGATTTATATGCTGATTTATGCGGATTTTCTTCCTTCCTTACTT 71
Db      5766 TTTAATATATTTGTTATTTGTTATTTTATATATATGTTATTTGTTGTTTTTTTACTT 5707

QY      72  TTATTTATTTCCCAATTTTCTTAAGCAATATTTCTTTGCTAATCAATAAATATCAAA 131
Db      5706 ATATTTCTATTTTATTTATATAAAATTAATAATTAATTAATAAATAAATAAATAA 5647

QY      132 AGAAAAAAAAGCTGAAACCAACGCTTGAAAAAGGAAA 169
Db      5646 AAAAAAAAATTTAAATTAATAAATAAATAAATAAATAA 5609

RESULT 15
DDP2PLAS
LOCUS
DEFINITION
Dictyostelium discoideum ddp2 plasmid DNA containing gene essential
for replication.
ACCESSION
X51478
VERSION
X51478.1 GI:7307
KEYWORDS
plasmid; plasmid Ddp2; rep gene.
SOURCE
Dictyostelium discoideum
ORGANISM
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE
1 (bases 1 to 5852)
AUTHORS
Slade,M.B., Chang,A.C. and Williams,K.L.
TITLE
The sequence and organization of Ddp2, a high-copy-number nuclear
plasmid of Dictyostelium discoideum
JOURNAL
Plasmid 24 (3), 195-207 (1990)
MEDLINE
91172902
PUBMED
2077544
REFERENCE
2 (bases 1 to 5852)
AUTHORS
Chang,A.C., Slade,M.B. and Williams,K.L.
TITLE
Identification of the origin of replication of the eukaryote
Dictyostelium discoideum nuclear plasmid Ddp2
JOURNAL
Plasmid 24 (3), 208-217 (1990)
MEDLINE
91172903
PUBMED
2077545
REFERENCE
3 (bases 1 to 5852)
AUTHORS
Slade,M.B.
TITLE
Direct Submission
JOURNAL
Submitted (24-JAN-1990) Slade M.B., School of Biological Sciences,
Macquarie University, Sydney NSW 2109, Australia
FEATURES
     source
       1..5852
         /organism="Dictyostelium discoideum"
         /mol_type="genomic DNA"
         /strain="WS380B"
         /db_xref="taxon:44689"
         /cell_line="HU2362"
         572..1072
           /rpt_type=INVERTED
           915..924
             /note="palindromic sequence tgtcatgaca"
           938..947
             /note="palindromic sequence tgtcatgaca"
           1100..1109
             /note="palindromic sequence tgtcatgaca"
           1137..1638
             /rpt_type=INVERTED
             1263..1272
               /note="palindromic sequence tgtcatgaca"
               1285..1294
                 /note="palindromic sequence tgtcatgaca"
                 1643..1652
                   /note="palindromic sequence tgtcatgaca"
                   1665..1674
                     /note="palindromic sequence tgtcatgaca"
                     2172..2181
                       /note="palindromic sequence tgtcatgaca"

```

```
/notes="palindromic sequence tggcatgaca"
2378. .5041
/genes="rep"
CDS
2378. .5041
/genes="rep"
/notes="essential for plasmid replication"
/codon_start=1
/protein_id="CAA35843.1"
/db_xref="GI:7308"
/db_xref="SPTREMBL:Q23913"
/translation="MDLISWDRPFKFFVILLLEPKGCKRNDVRLSDVLDIISGIYSP
RTFVLKEVRAVAVSYDESEIDLFRLGSVFPGTSLYSYIPGIFSLKDFLLISKTSGK
IRVSDVQALLIFDHFPSRISDKQVFRKDIIPGRTFPEKSISEYKISDGRAAGVSWFN
LVSKITCYKNHPLFAENPTYKHVDIFISMLSLVHGIIIVDSQNEDENNVAMYSNLPFV
DLEKSDIPGAVQSRVTNTRGRSNNLNPTTTTTTTTTTTTTTTTTTTTTTTTTTRKRSDD
SVQEQSSQPKTSRSGSLKDVRIINNIIVSDSSSDSDVMSVSNRLKCVLLKAVVNGK
EGLEVVKEVLQDQKNYATGLLENIIPHNSKSERVITLSSFFFEIASKINDYEVKFS
ELSIDVESAKRLTFEKNNTNLIIPNNFKEGFEFLVPIVNGIASIVFVSPNNYSSG
SPANVESALKLHLICISLNGINGFUSIRSIITFDTKSIITKDLIPMSKRMLEDQGFRR
LRDANNNSKKSVQSDISGIDETDKLISFVHEFINDLNLYKLSKEEDGLMLVDFF
TSTLFRNPNMSIDNKVGMFHCBSISKFQSCNHSIDNLVLSFTNNIKNISODNE
NELKKYSIMVSDFRNVKVPKPISEPKRTIITFTNNSYNANRVAFDDISGSI
ITNVKNIHAKGORNFIETLLGSTRIRAFPCAPCLIQINNFKATDKLIDDQSVNH
QTASLEIKNLSYLPDLIKVGRSTVGTIKGETAPIIINSEETFFSISCLDIRFSALI
SKTKLSQLPFTAPDERYNKTNILKVLQDCBELTFTFLNNYKIANKLSTIENLYNNF
MGLEDEDEDEDEDEDEDEDEDEG"
BASE COUNT 2296 a 652 c 709 g 2195 t
ORIGIN
Query Match 9.3%; Score 49.2; DB 3; Length 5852;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 12 TGTAGTCGGCTGATTATATCTGATTATGCGGATTTGCTTCTCTTTTACTT 71
Db 5767 TTTAATATTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTACTT 5708
QY 72 TTATTTATTTCCAAATTTTCTTAAGCAATATTTCTTCTTAATCAATAATATCAAA 131
Db 5707 ATATTTCTATTTTATTTATATAAATTAATAATTAATAATAAATAAATAA 5648
QY 132 AGAAAAAACTGAAGCAACGCTTGAAGAAAGGAAA 169
Db 5647 AAAAAAAATTTAAATTAATAAATAAATAAATAAATAA 5610
RESULT 16
A21622/c
LOCUS Dictoselium plasmid Ddp2 Rep gene. linear PAT 07-JUL-1994
DEFINITION Dictoselium plasmid Ddp2 Rep gene.
ACCESSION A21622
VERSION A21622.1 GI:583618
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM Dictoselium
REFERENCE 1 (bases 1 to 5852)
AUTHORS Noegel,A.A.
TITLE Sequence and analysis of chromosome 2 of Dictoselium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
FEATURES
source 1. .5852
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
2378. .5041
/genes="REP gene from a recombinant Ddp-2 like plasmid in Dictyostelium spp"
2378. .5041
/genes="REP gene from a recombinant Ddp-2 like plasmid in Dictyostelium spp"
/codon_start=1
/notes="palindromic sequence tggcatgaca"
2378. .5041
/genes="rep"
CDS
2378. .5041
/genes="rep"
/notes="essential for plasmid replication"
/codon_start=1
/protein_id="CAA35843.1"
/db_xref="GI:7308"
/db_xref="SPTREMBL:Q23913"
/translation="MDLISWDRPFKFFVILLLEPKGCKRNDVRLSDVLDIISGIYSP
RTFVLKEVRAVAVSYDESEIDLFRLGSVFPGTSLYSYIPGIFSLKDFLLISKTSGK
IRVSDVQALLIFDHFPSRISDKQVFRKDIIPGRTFPEKSISEYKISDGRAAGVSWFN
LVSKITCYKNHPLFAENPTYKHVDIFISMLSLVHGIIIVDSQNEDENNVAMYSNLPFV
DLEKSDIPGAVQSRVTNTRGRSNNLNPTTTTTTTTTTTTTTTTTTTTTTTTTTRKRSDD
SVQEQSSQPKTSRSGSLKDVRIINNIIVSDSSSDSDVMSVSNRLKCVLLKAVVNGK
EGLEVVKEVLQDQKNYATGLLENIIPHNSKSERVITLSSFFFEIASKINDYEVKFS
ELSIDVESAKRLTFEKNNTNLIIPNNFKEGFEFLVPIVNGIASIVFVSPNNYSSG
SPANVESALKLHLICISLNGINGFUSIRSIITFDTKSIITKDLIPMSKRMLEDQGFRR
LRDANNNSKKSVQSDISGIDETDKLISFVHEFINDLNLYKLSKEEDGLMLVDFF
TSTLFRNPNMSIDNKVGMFHCBSISKFQSCNHSIDNLVLSFTNNIKNISODNE
NELKKYSIMVSDFRNVKVPKPISEPKRTIITFTNNSYNANRVAFDDISGSI
ITNVKNIHAKGORNFIETLLGSTRIRAFPCAPCLIQINNFKATDKLIDDQSVNH
QTASLEIKNLSYLPDLIKVGRSTVGTIKGETAPIIINSEETFFSISCLDIRFSALI
SKTKLSQLPFTAPDERYNKTNILKVLQDCBELTFTFLNNYKIANKLSTIENLYNNF
MGLEDEDEDEDEDEDEDEDEDEG"
BASE COUNT 2296 a 652 c 709 g 2195 t
ORIGIN
Query Match 9.3%; Score 49.2; DB 6; Length 5852;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 12 TGTAGTCGGCTGATTATATCTGATTATGCGGATTTGCTTCTCTTTTACTT 71
Db 5767 TTTAATATTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTACTT 5708
QY 72 TTATTTATTTCCAAATTTTCTTAAGCAATATTTCTTCTTAATCAATAATATCAAA 131
Db 5707 ATATTTCTATTTTATTTATATAAATTAATAATTAATAATAAATAAATAA 5648
QY 132 AGAAAAAACTGAAGCAACGCTTGAAGAAAGGAAA 169
Db 5647 AAAAAAAATTTAAATTAATAAATAAATAAATAAATAA 5610
RESULT 16
A21622/c
LOCUS Dictoselium plasmid Ddp2 Rep gene. linear PAT 07-JUL-1994
DEFINITION Dictoselium plasmid Ddp2 Rep gene.
ACCESSION A21622
VERSION A21622.1 GI:583618
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM Dictoselium
REFERENCE 1 (bases 1 to 5852)
AUTHORS Noegel,A.A.
TITLE Sequence and analysis of chromosome 2 of Dictoselium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
FEATURES
source 1. .5852
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
2378. .5041
/genes="REP gene from a recombinant Ddp-2 like plasmid in Dictyostelium spp"
2378. .5041
/genes="REP gene from a recombinant Ddp-2 like plasmid in Dictyostelium spp"
/codon_start=1
/notes="palindromic sequence tggcatgaca"
2378. .5041
/genes="rep"
CDS
2378. .5041
/genes="rep"
/notes="essential for plasmid replication"
/codon_start=1
/protein_id="CAA35843.1"
/db_xref="GI:7308"
/db_xref="SPTREMBL:Q23913"
/translation="MDLISWDRPFKFFVILLLEPKGCKRNDVRLSDVLDIISGIYSP
RTFVLKEVRAVAVSYDESEIDLFRLGSVFPGTSLYSYIPGIFSLKDFLLISKTSGK
IRVSDVQALLIFDHFPSRISDKQVFRKDIIPGRTFPEKSISEYKISDGRAAGVSWFN
LVSKITCYKNHPLFAENPTYKHVDIFISMLSLVHGIIIVDSQNEDENNVAMYSNLPFV
DLEKSDIPGAVQSRVTNTRGRSNNLNPTTTTTTTTTTTTTTTTTTTTTTTTTTRKRSDD
SVQEQSSQPKTSRSGSLKDVRIINNIIVSDSSSDSDVMSVSNRLKCVLLKAVVNGK
EGLEVVKEVLQDQKNYATGLLENIIPHNSKSERVITLSSFFFEIASKINDYEVKFS
ELSIDVESAKRLTFEKNNTNLIIPNNFKEGFEFLVPIVNGIASIVFVSPNNYSSG
SPANVESALKLHLICISLNGINGFUSIRSIITFDTKSIITKDLIPMSKRMLEDQGFRR
LRDANNNSKKSVQSDISGIDETDKLISFVHEFINDLNLYKLSKEEDGLMLVDFF
TSTLFRNPNMSIDNKVGMFHCBSISKFQSCNHSIDNLVLSFTNNIKNISODNE
NELKKYSIMVSDFRNVKVPKPISEPKRTIITFTNNSYNANRVAFDDISGSI
ITNVKNIHAKGORNFIETLLGSTRIRAFPCAPCLIQINNFKATDKLIDDQSVNH
QTASLEIKNLSYLPDLIKVGRSTVGTIKGETAPIIINSEETFFSISCLDIRFSALI
SKTKLSQLPFTAPDERYNKTNILKVLQDCBELTFTFLNNYKIANKLSTIENLYNNF
MGLEDEDEDEDEDEDEDEDEDEG"
BASE COUNT 2296 a 652 c 709 g 2195 t
ORIGIN
Query Match 9.3%; Score 49.2; DB 6; Length 5852;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 12 TGTAGTCGGCTGATTATATCTGATTATGCGGATTTGCTTCTCTTTTACTT 71
Db 5767 TTTAATATTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTACTT 5708
QY 72 TTATTTATTTCCAAATTTTCTTAAGCAATATTTCTTCTTAATCAATAATATCAAA 131
Db 5707 ATATTTCTATTTTATTTATATAAATTAATAATTAATAATAAATAAATAA 5648
QY 132 AGAAAAAACTGAAGCAACGCTTGAAGAAAGGAAA 169
Db 5647 AAAAAAAATTTAAATTAATAAATAAATAAATAAATAA 5610
RESULT 16
A21622/c
LOCUS Dictoselium plasmid Ddp2 Rep gene. linear INV 27-MAR-2003
DEFINITION Dictoselium discoideum chromosome 2 map 5515173-5817331 strain AX4, complete sequence.
ACCESSION A216977
VERSION A216977.2 GI:28828573
KEYWORDS HTG.
SOURCE Dictoselium discoideum
ORGANISM Dictoselium discoideum
REFERENCE 1 (bases 1 to 302156)
AUTHORS Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P., Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
FEATURES
MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
REFERENCE 2 (bases 1 to 302156)
AUTHORS Baumgart,C.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 302156)
AUTHORS Baumgart,C.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 302156)
AUTHORS Baumgart,C.
TITLE Direct Submission
```


[illegible]


```

QY      146 AAGCAACGCTTGAAGAAAGGAAGTTAGCCCTA 179
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      35400 AAAAAAATCGAAAAATGAAATTTACATTA 35367

RESULT 22
AC091136/c
LOCUS   AC091136.2 GI:17047133
DEFINITION Homo sapiens chromosome 11 clone RP5-873F21 map 11, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC091136
VERSION   HTG; HTGS PHASE0.
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 39989)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
          Barna,N., Bastien,V., Boguslavskiy,L., Bouckhalter,B., Brown,A.,
          Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
          Collins,S., Collumore,A., Cooke,P., Dearellano,K., Dewar,K.,
          Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
          Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
          Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
          MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
          McKernan,K., McPheeters,R., Meldrim,J., Meneus,I.,
          Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
          Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
          Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
          Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
          Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
          Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
          Theodore,J., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
          Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
          Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
          Direct Submission
          Submitted (01-APR-2001) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          On Nov 22, 2001 this sequence version replaced gi:13493100.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          ----- Genome Center
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WIBR
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence_submissions@genome.wi.mit.edu
          ----- Project Information
          Center project name: L12134
          Center clone name: 873_F_21
          -----
          * NOTE: This record contains 38 individual
          * sequencing reads that have not been assembled into
          * contigs. Runs of N are used to separate the reads
          * and the order in which they appear is completely
          * arbitrary. Low-pass sequence sampling is useful for
          * identifying clones that may be gene-rich and allows
          * overlap relationships among clones to be deduced.
          * However, it should not be assumed that this clone
          * will be sequenced to completion. In the event that
          * the record is updated, the accession number will
          * be preserved.
          * 1
          * 939 1038: contig of 938 bp in length
          * 939 1038: gap of 100 bp
          * 1039 2183: contig of 1145 bp in length

```

FEATURES

Location/Qualifiers

```

source
1..39989
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone_lib="RP11 Human PAC library 5"
/clone_id="8024 c 7083 g 10567 t 3932 others"

BASE COUNT 10383 a 8024 c 7083 g 10567 t 3932 others
ORIGIN

Query Match
Best Local Similarity 9.1%; Score 48.2; DB 2; Length 39989;
Matches 127; Conservative 1; Mismatches 114; Indels 1; Gaps 1;

QY 26 ATTATATGCTGATTTATGGGTGATTTGCTTCCTTCTTTATACATTTTATTTATTTATTTCCCAA 85
Db 39386 ATCATATACCTGATATGAATGATTTTATTTTCCTTGTATATGATTTCTCTATTTCCCAA 39327

QY 96 ATTTTCTTAAGCAATATTT-CTTTGCTAATCAATAATATCAAAAGAAAAAAACT 144
Db 39326 CATTATGCAATAACATATTTATTTTATTTAAATAGTATCATAGAAAAAGGGAA 39267

QY 145 GAAACCAAGCTTGAAAAAGAAAGTTAGCCCTATCGGTATATTTTGGAAAGTTGTAAA 204
Db 39266 AACGTCACATACATTGAGTATTATGACACAAAATTTCTGTATTAGATTGCTACTCTCCAC 39207

QY 205 ATACTACRTGTTCTCTTCTAAGTCCACCTCTCTGTTTCTTTGACGAGGAAGAAG 264
Db 39206 AATACAGAGCTTCATTTTACATGACAGACCCCTTTATCTGAGATAGAAAGGGGAAA 39147

QY 265 CAG 267
Db 39146 AAG 39144

RESULT 23
AC021298/c
LOCUS 146150 bp DNA linear HTG 06-AUG-2002
DEFINITION Homo sapiens chromosome 11 clone RP11-222M17 map 11, WORKING DRAFT
SEQUENCE, 5 ordered pieces.
ACCESSION AC021298
VERSION AC021298.5 GI:22123002
KEYWORDS HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 146150)
Birren,B., Nusbaum,C. and Lander,E.
Unpublished
Homo sapiens chromosome 11, clone RP11-222M17
2 (bases 1 to 146150)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fensetor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,G., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Raymond,C., Riley,R., Rothman,D.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riese,C., Rogov,P.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo.A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

```

REFERENCE
AUTHORS
3 (bases 1 to 146150)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Baatien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Canarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Miengo,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Riese,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:12740251.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4206
Center clone name: 222.M.17
----- Summary Statistics
Sequencing vector: M13; M77815; 4% of reads
Sequencing vector: Plasmid; n/a; 96% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 14493 bases at least Q40
Consensus quality: 145263 bases at least Q30
Consensus quality: 145559 bases at least Q20
Insert size: 140000; agarose-fp
Quality coverage: 8.6 in Q20 bases; agarose-fp
Quality coverage: 8.2 in Q20 ba
* NOTE: this is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 49523: contig of 49523 bp in length
* 49524 49623: gap of 100 bp
* 49624 51342: contig of 1719 bp in length
* 51343 51442: gap of 100 bp
* 51443 55402: contig of 3960 bp in length
* 55403 55502: gap of 100 bp
* 55503 93446: contig of 37944 bp in length
* 93447 93546: gap of 100 bp
* 93547 146150: contig of 52604 bp in length.
* Location/Qualifiers
1. 146150
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone_lib="RP11-222M17"
/clone_lib="RP11-11 Human Male BAC"
TITLE
JOURNAL

```



```

/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-624D11"
/clone_lib="RPC1-11 Human Male BAC"
997. .917
/rpt_family="AT_rich"
1150. .1169
/rpt_family="(T)n"
complement(1820. .1987)
/rpt_family="MIR"
complement(2145. .2206)
/rpt_family="MIR"
3207. .3495
/rpt_family="L2"
3995. .4080
/rpt_family="MIR3"
4157. .4648
/rpt_family="MLT1E"
4894. .4958
/rpt_family="MIR3"
4962. .5018
/rpt_family="MERSA"
5019. .5783
/rpt_family="L1P4"
5784. .5903
/rpt_family="MERSA"
complement(5956. .6069)
/rpt_family="MIR"
complement(6952. .7163)
/rpt_family="MIR"
complement(7264. .7478)
/rpt_family="MLT1-int"
7479. .7761
/rpt_family="AluSc"
complement(7762. .8046)
/rpt_family="MLT1-int"
complement(8103. .8751)
/rpt_family="MLT1-int"
complement(8770. .9068)
/rpt_family="MLT2C1"
9069. .9368
/rpt_family="AluSx"
complement(9369. .9480)
/rpt_family="MLT2C1"
9634. .9925
/rpt_family="AluJo"
9997. .10128
/rpt_family="AluSp/q"
complement(11650. .11717)
/rpt_family="LTR33"
11726. .12087
/rpt_family="MLT1A2"
12577. .12598
/rpt_family="(CAAA)n"
12688. .12777
/rpt_family="Charlie8"
complement(12928. .13032)
/rpt_family="MIR"
13088. .13466
/rpt_family="L1M4"
15240. .15264
/rpt_family="AT_rich"
complement(15331. .15556)
/rpt_family="MIR"
complement(15708. .15842)
/rpt_family="L2"
18201. .18387
/rpt_family="MER117"
complement(18507. .18727)
/rpt_family="MIR"
complement(19164. .19252)
/rpt_family="L2"

repeat_region complement(21582. .21799)
/rpt_family="MIR"
repeat_region complement(22125. .22227)
/rpt_family="L1M4"
repeat_region 22264. .22287
/rpt_family="AT_rich"
repeat_region complement(22347. .22466)
/rpt_family="LTR16C"
repeat_region complement(22698. .22946)
/rpt_family="L1M4"
repeat_region 22947. .23343
/rpt_family="MLT1J1"
repeat_region 24245. .24502
/rpt_family="MIR"
repeat_region 25311. .25454
/rpt_family="L3"
repeat_region complement(25481. .25513)
/rpt_family="MERSA"
repeat_region 25514. .25533
/rpt_family="(CAAA)n"
repeat_region complement(25534. .25598)
/rpt_family="MERSA"
repeat_region complement(25599. .26466)

Query Match 9.1%; Score 48.2; DB 9; Length 155985;
Best Local Similarity 52.3%; Pred. No. 1.8;
Matches 127; Conservative 1; Mismatches 114; Indels 1; Gaps 1;

QY 26 ATTTATATGCTGATTATGGGTGATTTTGGTCTCTCTTATATCTTTATTTATTTATTTCCCAA 85
DB 57319 ATCACTATACCTGATTATGAATGATTTTATTTCTTTGTTATATATGATTTCTCTATTTCCAAA 57260
QY 86 ATTTTCTTAAAGCAATATTT-CTTGTCTATCAATAAATTTATCAAAAGAAAAAAACT 144
DB 57259 CATTATGCATTAACATATTTTACTTTTATTAATATAGTATCATAGAAAAAAGGAA 57200
QY 145 GAAAGCAACGCTTGAAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAAA 204
DB 57199 AACGTCACATCATGAGTATTATGACACAAAATTTCTGTATTTAGATTGCTACTCTCCAC 57140
QY 205 ATACTACRTGTTCTTCTTAAGTCCCACTCTCTGTTTCTTTGAGCAGGAAGAAG 264
DB 57139 AAAACAGAGCTTCATTTTACAATGACAGACCCCTTTATCTCTGATAGGAAGGGGAAA 57080
QY 265 CAG 267
DB 57079 AAG 57077

RESULT 25
AC068749
LOCUS AC068749 181996 bp DNA linear HTG 04-MAY-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-802012 map 11, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AC068749
VERSION AC068749.2 GI:8389606
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 181996)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
TITLE Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, P.,
JOURNAL Boguslavskiy, L., Boukhgeltser, B., Brown, A., Burkett, G.,
REFERENCE Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
AUTHORS Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

```

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,J., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meidrim,J., Meneus,L., Mhova,T., Miranda,C., Mienga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (08-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 9, 2000 this sequence version replaced gi:7717156.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L9955

Center clone name: 802_O12

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 171741 bases at least Q40

Consensus quality: 176943 bases at least Q30

Consensus quality: 179052 bases at least Q20

Insert size: 182000; agarose-fp

Quality coverage: 4.5 in Q20 bases; agarose-fp

Quality coverage: 4.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1109: contig of 1109 bp in length

* 1110 1209: gap of 100 bp

* 1210 2970: contig of 1761 bp in length

* 2971 3070: gap of 100 bp

* 3071 8242: contig of 5172 bp in length

* 8243 8342: gap of 100 bp

* 8343 14456: contig of 6114 bp in length

* 14457 14556: gap of 100 bp

* 14557 19217: contig of 4661 bp in length

* 19218 19317: gap of 100 bp

* 19318 24451: contig of 5134 bp in length

* 24452 24551: gap of 100 bp

* 24552 30373: contig of 5822 bp in length

* 30374 30473: gap of 100 bp

* 30474 38117: contig of 7644 bp in length

* 38118 38217: gap of 100 bp

* 38218 45565: contig of 7348 bp in length

* 45566 45666: gap of 100 bp

* 45666 57699: contig of 12034 bp in length

* 57700 57799: gap of 100 bp

* 57800 70834: contig of 13035 bp in length

* 70835 70935: gap of 100 bp

* 70935 84875: contig of 13941 bp in length

* 84876 84975: gap of 100 bp

* 84976 98366: contig of 13391 bp in length

* 98367 98466: gap of 100 bp

* 98467 117242: contig of 18776 bp in length

* 117243 117342: gap of 100 bp

* 117343 147820: contig of 30478 bp in length

* 147821 147920: gap of 100 bp

* 147921 181996: contig of 34076 bp in length.

FEATURES

source

1. .181996

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-802012"

/clone.lib="RPC1-11 Human Male BAC"

1. .1109

/note="assembly_fragment"

1210. .2970

/note="assembly_fragment"

3071. .8242

/note="assembly_fragment"

8343. .14456

/note="assembly_fragment"

14557. .19217

/note="assembly_fragment"

clone_end:77

vector_side:right

19318. .24451

/note="assembly_fragment"

24552. .30373

/note="assembly_fragment"

30474. .38117

/note="assembly_fragment"

38218. .45565

/note="assembly_fragment"

45666. .57699

/note="assembly_fragment"

57800. .70834

/note="assembly_fragment"

70935. .84875

/note="assembly_fragment"

clone_end:SP6

vector_side:left

84976. .98366

/note="assembly_fragment"

98467. .117242

/note="assembly_fragment"

117343. .147820

/note="assembly_fragment"

147921. .181996

/note="assembly_fragment"

BASE COUNT 55224 a 33965 c 35178 g 56119 t 1510 others

ORIGIN

Query Match 9.1%; Score 48.2; DB 2; Length 181996;
 Best Local Similarity 52.3%; Pred. No. 1.8;
 Matches 127; Conservative 1; Mismatches 114; Indels 1; Gaps 1;

QY 26 ATTATATGCTGATTTATGGGTGATTTGCTCTCTTTATACCTTTATTTATTTTCCCAA 85

DB 9746 ATCACTATACCTGATTTATGAATGATTTTATTTCTTTTGTATGATTTCTCTATTTCCAAA 9805

QY 86 ATTTTCTTAAGCAATATTTT-CTTTGCTTAATCAATAATTTATCAAAAGAAAAAAACT 144

DB 9806 CATTAAGCAATAACATATTTACTTTTATTAATAATGATCATAGAAAAAAGGAA 9865

QY 145 GAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGGAGTTGTAAA 204

DB 9866 AACGTCATCATATTGAGTATTATGACACAAAAATTTCTGTTATTAGATTGCTACTCTCAC 9925

QY 205 ATACTAGTGTCTCTTCTTAAGTCCCACTCTCTGTTTCTTTTGGAGGAAGAAG 264


```

Db      9926 AAAACAGAGCTTCATTATTTACATGACAGACCCCTTTATCTGAGATAGAGAAAGGGGAAA 9985
QY      265 CAG 267
        ||
Db      9986 AAG 9988

RESULT 26
AL356371
LOCUS   AL356371                256221 bp    DNA    linear    HTG 20-OCT-2001
DEFINITION Homo sapiens chromosome 11 clone RP1-135I9, *** SEQUENCING IN
PROGRESS ***, 22 unordered pieces.
ACCESSION AL356371
VERSION   AL356371.18 GI:14575185
KEYWORDS  HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
          Mclay, K.
          Direct Submission
          Submitted (16-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
          CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
          requests: clonerequest@sanger.ac.uk
          On Jun 28, 2001 this sequence version replaced gi:14529810.
COMMENT   ----- Genome Center
          Center: Sanger Centre
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: humquery@sanger.ac.uk
          ----- Project Information
          Center project name: d1135I9
          ----- Summary Statistics
          Assembly program: XGAP4; version 4.5
          Sequencing vector: M13; M77815; 1% of reads
          Sequencing method: plasmid; L08752; 98% of reads
          Chemistry: Dye-terminator ET-amersham; 2% of reads Chemistry:
          Dye-terminator Big Dye; 97% of reads
          Consensus quality: 250889 bases at least Q40
          Consensus quality: 252160 bases at least Q30
          Consensus quality: 252999 bases at least Q20
          Insert size: 254121; sum-of-contigs
          Insert size: 98381; 1.2% error; agarose-fp
          Quality coverage: 9.35x in Q20 bases; sum-of-contigs Quality
          coverage: 25.82x in Q20 bases; agarose-fp
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 22 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          1 2923: contig of 2923 bp in length
          * 2924: gap of 100 bp
          * 3024: contig of 3307 bp in length
          * 6331: gap of 100 bp
          * 6431: contig of 7855 bp in length
          * 14286: gap of 100 bp
          * 14386: contig of 11980 bp in length
          * 26366: gap of 100 bp
          * 26466: contig of 11549 bp in length
          * 38015: gap of 100 bp
          * 38115: contig of 2920 bp in length
          * 41035: gap of 100 bp
          * 41135: contig of 11417 bp in length
          * 52552: gap of 100 bp
          * 52652: contig of 4140 bp in length
          * 56792: gap of 100 bp
          * 56892: contig of 8242 bp in length
          * 65134: gap of 100 bp

65234 67244: contig of 2011 bp in length
67344: gap of 100 bp
69441: contig of 2097 bp in length
69541: gap of 100 bp
69542 91678: contig of 22137 bp in length
91679: gap of 100 bp
91779 97058: contig of 5280 bp in length
97158: gap of 100 bp
97159 99624: contig of 2466 bp in length
99724: gap of 100 bp
99725 106782: contig of 7058 bp in length
106783 106882: gap of 100 bp
106883 112388: contig of 5506 bp in length
112389 112489: gap of 100 bp
112489 116275: contig of 3786 bp in length
116276 116374: gap of 100 bp
116375 118939: contig of 2565 bp in length
118940 129647: contig of 10608 bp in length
129648 129747: gap of 100 bp
129748 136936: contig of 7189 bp in length
136937 137036: gap of 100 bp
137037 248507: contig of 11471 bp in length
248508 248607: gap of 100 bp
248608 256221: contig of 7614 bp in length.

FEATURES             Location/Qualifiers
     source            1..256221
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="RP1-135I9"
                     /clone_lib="RPC1-1"
     misc_feature       1..2923
                     /note="assembly_fragment:00208"
     misc_feature       3024..6330
                     /note="assembly_fragment:00427"
     misc_feature       6431..14285
                     /note="assembly_fragment:00726"
     misc_feature       14386..26365
                     /note="assembly_fragment:00955.0"
     misc_feature       26466..38014
                     /note="assembly_fragment:01016"
     misc_feature       38115..41034
                     /note="assembly_fragment:01599.0"
     misc_feature       41135..52551
                     /note="assembly_fragment:02394"
     misc_feature       52652..56791
                     /note="assembly_fragment:02491.0"
     misc_feature       56892..65133
                     /note="assembly_fragment:02823"
     misc_feature       65234..67244
                     /note="assembly_fragment:03451"
     misc_feature       67345..69441
                     /note="assembly_fragment:03517"
     misc_feature       69542..91678
                     /note="assembly_fragment:03857"
     misc_feature       91779..97058
                     /note="assembly_fragment:04538"
     misc_feature       97159..99624
                     /note="assembly_fragment:04662"
     misc_feature       99725..106782
                     /note="assembly_fragment:04852"
     misc_feature       108883..112388
                     /note="assembly_fragment:04898"
     misc_feature       112489..116274
                     /note="assembly_fragment:05674.0"
     misc_feature       116375..118939
                     /note="assembly_fragment:06302"
     misc_feature       119040..129647
                     /note="assembly_fragment:06497"
     misc_feature       129748..136936
                     /note="assembly_fragment:06540"

```



```
/note="assembly_fragment
clone end:T7
vector side:left"
misc_feature 85817..119636
/note="assembly_fragment"
misc_feature 119737..152336
/note="assembly_fragment"
BASE COUNT 51773 a 29398 c 27604 g 42706 t 855 others
ORIGIN
Query Match 9.0%; Score 47.8; DB 2; Length 152336;
Best Local Similarity 63.8%; Pred. No. 2.2;
Matches 90; Conservative 0; Mismatches 47; Indels 4; Gaps 1;
QY 26 ATTATATGCTGATTATGCTGATTGCTTCTTCTTATATCTTTTATTATTCCTCAA 85
DB 106738 ATTACATCTGTAATATGATGATTTCTTCTCTTAGATTTTGGATTTCTCA 106797
QY 86 ATTTTCTCT---TAAGCAATATTTCTTTGCTTAATCAATAATATCAAGAAAAAAA 141
DB 106798 AATTTTCTATCTGAGTATGATTATTTGTTATTAATCAGGAGTAGCAATTCAAAAAAGA 106857
QY 142 ACTGAAGCAACGCTTGAAAA 162
DB 106858 AAGAAAAATCTAGCTTTAATAA 106878
```

```
RESULT 30
AC079354
LOCUS AC079354 162554 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-1145F21 from 2, complete sequence.
ACCESSION AC079354
VERSION AC079354.4 GI:18098550
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 162554)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 162554)
AUTHORS Armstrong,J., Cotton,M., Hawkins,M. and Dignan,G.
TITLE The sequence of Homo sapiens BAC clone RP11-1145F21
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 162554)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 162554)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 9, 2002 this sequence version replaced gi:14550326.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
-----
Center project name: H_NH1145F21
-----
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frenken,B., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-107N15, 2000 bp overlap; the clone sequenced to the right is RP11-686O6. Actual start of this clone is at base position 1 of RP11-1145F21; actual end is at base position 162554 of RP11-1145F21.

Data from AC069148 and AC064836 was used to finish this clone.

AC079354.

FEATURES

source

Location/Qualifiers

1..162554
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-1145F21"
/clone_lib="RPCI-11"
329..433
/rpt_family="MIR"
1453..1493
/rpt_family="(CA)n"
1658..1775
/rpt_family="MIR"
1844..1916
/rpt_family="LI"
2232..2402
/note="similar to EST BB614631 (NID:G15395798)"
2583..2603
/rpt_family="AT_rich"
3403..3542
/rpt_family="LI"
3568..3595
/rpt_family="AT_rich"
3607..4123
/rpt_family="ERV1"
4423..4588
/rpt_family="(GGA)n"
4608..5023
/rpt_family="MaLR"
5049..5224
/note="similar to EST BB614631 (NID:G15395798)"
5692..5991
/rpt_family="Alu"
5973..6004
/rpt_family="AT_rich"
6261..6476
/rpt_family="MIR"
7016..7295


```
FEATURES
source
    true right end of clone RP11-317C20 is at 56073 in this sequence.
    Location/Qualifiers
    1. .182012
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="9"
    /map="q31.1-31.3"
    /clone_lib="RPCI-11.2"
    /clone="RP11-413C10"
    2456. .2511
    /notes="28 copies 2 mer aa 71% conserved"
    2622. .2787
    /notes="MIR repeat: matches 20. .196 of consensus"
    2907. .3078
    /notes="L2 repeat: matches 2566. .2749 of consensus"
    3110. .3345
    /notes="MIR repeat: matches 21. .237 of consensus"
    3927. .4102
    /notes="88 copies 2 mer tt 57% conserved"
    4108. .4406
    /notes="AluX repeat: matches 1. .303 of consensus"
    6277. .6330
    /notes="27 copies 2 mer ta 72% conserved"
    7275. .7467
    /notes="MIR repeat: matches 10. .219 of consensus"
    7879. .7995
    /notes="LIPAS repeat: matches 6029. .6145 of consensus"
    9391. .9426
    /notes="18 copies 2 mer tc 94% conserved"
    9525. .9811
    /notes="MER47A repeat: matches 46. .333 of consensus"
    10892. .10959
    /notes="L1M1 repeat: matches 718. .785 of consensus"
    10998. .11075
    /notes="3 copies 26 mer 92% conserved"
    11111. .11148
    /notes="19 copies 2 mer ac 92% conserved"
    11507. .12390
    /notes="LIPAS repeat: matches 5254. .6143 of consensus"
    15656. .16353
    /notes="L1M3 repeat: matches 5397. .6149 of consensus"
    16387. .16491
    /notes="FLAM A repeat: matches 22. .127 of consensus"
    16557. .16682
    /notes="L1M3 repeat: matches 5189. .5329 of consensus"
    17439. .17610
    /notes="MIR repeat: matches 49. .241 of consensus"
    19584. .20420
    /notes="LIPAS repeat: matches 5308. .6143 of consensus"
    20813. .21163
    /notes="THB1A repeat: matches 1. .354 of consensus"
    21164. .22730
    /notes="THB1A-internal repeat: matches 1. .1578 of consensus"
    22735. .23078
    /notes="THB1A repeat: matches 3. .354 of consensus"
    23719. .24027
    /notes="AluSq repeat: matches 1. .311 of consensus"
    25619. .25765
    /notes="MER5A repeat: matches 15. .187 of consensus"
    25784. .25850
    /notes="MER5B repeat: matches 103. .173 of consensus"
    26130. .26262
    /notes="MER5A repeat: matches 7. .141 of consensus"
    26884. .27199
    /notes="AluSq repeat: matches 1. .312 of consensus"
    28112. .28361
    /notes="L2 repeat: matches 2243. .2487 of consensus"
    28389. .28691
    /notes="AluSC repeat: matches 1. .303 of consensus"
    28694. .28717
    /notes="HSMAR1 repeat: matches 1264. .1287 of consensus"
```

```
repeat_region
    29037. .29333
    /note="AluJb repeat: matches 1. .299 of consensus"
    29888. .30195
    /note="AluY repeat: matches 1. .308 of consensus"
    30289. .30601
    /note="AluSc repeat: matches 1. .311 of consensus"
    30778. .30937
    /note="MER52A repeat: matches 1597. .1755 of consensus"
    30940. .31306
    /note="MER52A repeat: matches 1096. .1470 of consensus"
    31307. .31608
    /note="AluSq repeat: matches 1. .303 of consensus"
    31609. .31902
    /note="MER52A repeat: matches 2. .94 of consensus"
    31974. .32054
    /note="L2 repeat: matches 2606. .2691 of consensus"
    32199. .32315
    /note="MER5B repeat: matches 60. .177 of consensus"
    32401. .32488
    /note="L2 repeat: matches 2452. .2550 of consensus"
    32903. .33206
    /note="L1M2 repeat: matches 517. .221 of consensus"
    33241. .34061
    /note="L1M3e repeat: matches 455. .368 of consensus"
    34739. .35429
    /note="L1M2 repeat: matches 725. .1400 of consensus"
    35430. .35735
    /note="AluVa5 repeat: matches 1. .306 of consensus"
    35736. .35873
    /note="L1M2 repeat: matches 1400. .1539 of consensus"
    35894. .36856
    /note="L1M4 repeat: matches 3844. .4859 of consensus"
    36857. .37311
    /note="L1M2 repeat: matches 5333. .5787 of consensus"
    37311. .38137
    /note="LIPAL3 repeat: matches 3313. .4197 of consensus"
    38139. .38833
    /note="LTR12 repeat: matches 1. .664 of consensus"
    38841. .39840
    /note="LIPAL3 repeat: matches 4195. .5182 of consensus"
    39843. .39876
    /note="17 copies 2 mer ta 82% conserved"
    39913. .40870
    /note="LIPAL3 repeat: matches 5190. .6156 of consensus"
    40873. .41084
    /note="L1M4 repeat: matches 5769. .5974 of consensus"
    41085. .41373
    /note="AluSc repeat: matches 1. .297 of consensus"
    41374. .41623
    /note="L1M4 repeat: matches 5974. .6300 of consensus"
    41634. .41856
    /note="L1M4c repeat: matches 1811. .1627 of consensus"
    41943. .42051
    /note="AluSq repeat: matches 1. .109 of consensus"
    42068. .42344
    /note="LIPAS repeat: matches 5861. .6142 of consensus"
    42345. .42553
    /note="AluSq repeat: matches 101. .310 of consensus"
    42688. .43082
    /note="L1M1 repeat: matches 1940. .2717 of consensus"
    43060. .43547
    /note="L1M9 repeat: matches 2434. .2924 of consensus"
    43555. .43753
    /note="AluJb repeat: matches 86. .285 of consensus"
    43754. .43987
    /note="117 copies 2 mer ta 76% conserved"
    43767. .44004
    /note="7 copies 34 mer 73% conserved"
    44065. .47403
    /note="L1M9 repeat: matches 2919. .6260 of consensus"
    48015. .48129
    /note="FLAM_C repeat: matches 30. .142 of consensus"
    48134. .48314
```

```

repeat_region /rpt_family="Alu"
7297. .7325
repeat_region /rpt_family="(CAA)n"
7485. .8131
repeat_region /rpt_family="L1"
10503. .10824
repeat_region /rpt_family="Alu"
10621. .10648
repeat_region /rpt_family="(A)n"
10795. .10824
repeat_region /rpt_family="(A)n"
10877. .10970
repeat_region /rpt_family="(TG)n"
11032. .11182
repeat_region /rpt_family="(TG)n"
11275. .12033
repeat_region /rpt_family="ERV1"
12196. .12323
repeat_region /rpt_family="MER1_type"
12907. .13340
misc_feature /note="similar to EST A1190728 (NID:G3741937) qd54h06.xl"
13399. .13534
repeat_region /rpt_family="MIR"
13952. .14017
repeat_region /rpt_family="L2"
14181. .14203
repeat_region /rpt_family="(TGAA)n"
14395. .14420
repeat_region /rpt_family="(GTTTG)n"
14461. .14538
repeat_region /rpt_family="Mariner"
15030. .15052
repeat_region /rpt_family="AT-rich"
15595. .15891
repeat_region /rpt_family="Alu"
15754. .15777
repeat_region /rpt_family="(T)n"
15962. .16507
repeat_region /rpt_family="MaLR"
16835. .16953
repeat_region /rpt_family="MIR"
17029. .17180
repeat_region /rpt_family="MER103"
17670. .17790
repeat_region /rpt_family="ERV1"
18303. .18577
repeat_region /rpt_family="Alu"
19009. .19102
repeat_region /rpt_family="Alu"
19103. .19470
repeat_region /rpt_family="MaLR"
19471. .19680
repeat_region /rpt_family="Alu"
19681. .19794
repeat_region /rpt_family="MIR"
20190. .20573
repeat_region /rpt_family="MaLR"
20880. .21004
repeat_region /rpt_family="MIR"
21787. .22090
repeat_region /rpt_family="Alu"
22975. .22999
repeat_region /rpt_family="AT-rich"
23112. .23153
repeat_region /rpt_family="CT-rich"
23196. .23498
repeat_region /rpt_family="Alu"
24077. .24377
repeat_region /rpt_family="Alu"
25224. .25259
repeat_region /rpt_family="(T)n"
25230. .25539
repeat_region /rpt_family="Alu"

```

```

repeat_region 25897. .25946
/rpt_family="(TGG)n"
repeat_region 25973. .26158
/rpt_family="MIR"
repeat_region 27003. .27077
/rpt_family="L2"
repeat_region 27086. .27294
/rpt_family="L1"
repeat_region 27550. .27857
/rpt_family="Alu"

```

```

Query Match          9.0%; Score 47.8; DB 9; Length 162554;
Best Local Similarity 55.0%; Pred. No. 2.2;
Matches 94; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 26 ATTATATGCTGANTTATGGGTGATTTTGGCTTCCTTATATCTTTATTTATTTCCAA 85
Db 82152 AGTTATGAGTTGACTGTGGGGGTTTGTCTCTTTATGATTTTGTGTGTTTCCAA 82211

QY 86 ATTTTCTTAAGCAATATTTCTTTGCTAAATCAATAAATTTATCAAAAGAAAAAATCTG 145
Db 82212 TTTTCTCAATAAATGTAICTTCCATPAATCAAAAATTTATTAAGTAATAAACA 82271

QY 146 AAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAA 196
Db 82272 GATGTAGACTAGTCCATATGAAGTCTTGGTCTGGAGACTCTTCATTCTAA 82322

```

```

RESULT 31
AL359846
LOCUS Human DNA sequence from clone RP11-413C10 on chromosome
DEFINITION 931.1-31.3, complete sequence.
ACCESSION AL359846
VERSION AL359846.11 GI:11611367
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 182012)
AUTHORS Babbage,A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11557924.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-413C10 is from the library RP11-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6

```

```

COMMENT
The left end of clone RP11-31J20 is at 123231 in this sequence. The

```



```
complement (13818..14342)
/note="ORF_ID:dd_02705"
/codon_start=1
/product="TRANSPORTER"
/protein_id="AAM44362.1"
/db_xref="GI:21240651"
/translation="MGEKSSQKTLKQNLKSKSGKICCEIVFPPIIIVLLVPAAIL
VLVQPKFNYDIKTQFSNRNENNIYGGKAGSLNVEQGVDMKQFQSLNENLK
SISVDVVFKEINDSEMEYFQINSQVGFQWESQOLSVANTTNTTNSFKY
STRLDNSKNGKNDGSDSVLTQWAVIQIADQAIQGYGLDYKLVINGQRY
PDYVELKQWNGRESVPKSGSVFSAALIFTFLAVLTVVEKETKREGMSIG
LNQCYFSTWITSVLPALPIDLIIVILKSGQVHSWIIIVTLILYLLTLQLLA
FIFSMFQKSGAGLTLFTLILINIGIFGEYEIINIHILKLCPCSPGIGACSF
YIMSIRDPSTTINLNOVISEKQIIGTFVFNIIFFYLWLDKIVKTEYGTKEP
YELFKRYWTVGAGIKKNGKNEYNDIESTIQNNNDNIEMVPIEVNRKTTISIR
NURKEFGDGLIARVNDLDMFQQLHGLGPNQSGKSTISMLTGLLSTGTAF
ILGNDITVQMNREIKCTGVCLOTILNQLTVELEIIVASLKGITKKQKIKESLKM
ANEIDLGKHLTPAGSLGGQKRLCLGAPIGRSTIPLDEVSGMDPSRRKXWDF
LUSHKGGKTLIITHYLDADYLSRISISHGKLTIDGSSFLKKNYGVGYLLTCSK
SNTIDQFNVDQVTFIRDOI PDVTYLSNAGSEISFRLPTASLPVFSKFPDFDENS
NPHIDSYGISVTVLEVFQIGTIDTDALSADGNDDELKRGISVSTGIRSKQ
LKALLIKRVKTSKDFKAPLLSLLPLAVIIGSIIVFKEVDNEVIFYNSTPLTFSL
LQOYGKNDIVPIQLAGTSEFPQNFPLSNVFNQFOYLNNSINNOYLNNYKOSGG
SNFTPIPLSSILNTTKTVISYNSLNFENYIHSMPVUNLINDALLNENHGIESTS
LFDHILTSFQASQGMQLQSIYFIVIMAGFSLMAGSFAGSISQERTNRKRLIYI
SCKKIIYVSNLLMDYFPAFILLITCIILAI VDDKNDQFLGLSLVLSLSIIP
LSYLSFKFSFGKSTGAITAHFSGIVMTIAMILRIEVIKSSSLNENIADIID
VFNILSPYAFGRVIFIIISGFGSLRGLKALIDDYWSLHYASTPIIILSVHVMTIF
ILLIDYSPFKGYLRNPKTLIPPPIDEDSDVSAERIRLSMSPIRSTDOGGGGGD
MLOYGLHLKFLTGKGNKNNAVYSTLIGPOQTFGLLGLGAGKTTVSMAGDIL
PTSGQHLHGLVTRDAQLRGVSCPODALITLLTAREQSLICAIGVPPDDKIS
LYVEAFIKMDLKGKIANSTGGYSGGNKRVKSLTAMLGNSVWLIDEASGDFIIR
YRQCVISLNGKNTIILTSMSSEIQAICSRMTIMRQCFKLASTQHKSKFGAGY
SYEVFKKSGLEGVQSLQCFPNATILDOHLMASLELPPNPNPNIKVSEIFN
ILSTELSSILDVSVSQTLEQVFLKLTGATHEDRLNLNQOQQOQTIPNSD"
complement (join (7575..8146,8235..8364))
/note="ORF_ID:dd_02702"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM44363.1"
/db_xref="GI:21240652"
/translation="MEIKYFLVLLVGLFVLSIVNYPKGVITINQPKINIYQLDV
NPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPN
HPHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
PHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPHPH
SYQGNRLIFIKRDY"
9039..10208
/note="ORF_ID:dd_02724"
/codon_start=1
/product="POTASSIUM CHANNEL REGULATORY FACTOR"
/protein_id="AAM44364.1"
/db_xref="GI:21240653"
/translation="MSSLNLKLEQLEKKEKEKEKEKEKENNNNNNNNNNNN
EPATKYTRGEIDKQLKQKEERIKEKEKEKEKENNNNNNNNNNNNNNNNN
NINIDNDSNKINKLNEGFLPDIVIQRLKRSQPIITGETSIDRAWRLKLE
ENDPIYQDGNDEFANILKQIQRKWKHQDADNEKEKENNDKKQOQQQENEE
SEEELEIRIDSKSKELIYFLKGLNDWITNLNENSDDEKSKRQGIABEATYQ
CASHIOPLTSTLSDILEINEIVEFCKSREYVRANDPELQMAIGNAPWPMGV
TWGHIHERSREKIRFNSOVARVLDNEMQRKYQAVKRLMTFCQTOYTPAPRSIG"
join (11612..11806,12045..12147,12280..12424,12526..12693,
12802..13164,13244..13607)
/note="ORF_ID:dd_02721"
/codon_start=1
/product="Putative Chromosome segregation ATPase"
/protein_id="AAM44365.1"
/db_xref="GI:21240654"
/translation="MNFRELDTSNGVKAESSFRTKVVLAREIPWNGPASSNITSEQ
YNLISKYDKTDABKEKAANSAYVFNFPNFTNENIETIYQLLTLINEIIEIDP
RAAGAFSTKDDKSYPSVFFRLNREDYATNLHASIAQIMCAGTPQNDVSEF
FWMLKLARNNSSEVEVGLAQSLLLKDDFRIFKFRKSGALLNLQALSTSLN
IQLLYETIYAIWLTLYKDI AAYSGTVANLVQLVTVAKEKIVRLSLTSLNLL
NGKSNEMDNGFVRLNLTINKWGDIDIPADIEVLINGLAKIDINMSFNYKTEI
ISGEWTFVPHKSEFNKISFENNYQIKHLGHILKTSQSTPLQUSIACHDICE
FVNHSGRGKAIMTITNQTRYHGYDVKSNEEVKNQALFALQRMMLNNWEYNLAK"
```

```
complement (13818..14342)
/note="ORF_ID:dd_02721"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM44366.1"
/db_xref="GI:21240655"
/translation="MNFENFKIIDEIKTEILLKNADFIHHIPIKHYSKLIHNNTDNNN
DNGSYENWRIIYKDCNNISIPMFIKPHLFIGNDCUKPINDKOLELLFQCYCFDID
LRNLIIINSYFNFTNSMISKAIDNRIDLYLKYLENGEIKNPNLSLQLHLKQNL
SKSYDNCFLDNLN"
complement (14672..17431)
/note="ORF_ID:dd_02711"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM44367.1"
/db_xref="GI:21240656"
/translation="MESPNVEILKWKPHNKYLFSLILKEIKKEMVNVVYLGVHVDN
RRRTTSMEMWNNNEQFELKUKINSNDIEILKKSISELPSINDPSLIEQLLTQ
LQGNKQFQIKINEMLPTLPIITNSLRFNNTNIFQLLSNLSUSSQFQIYSTL
LECAIEFLQPNLTSLINSPNLTQDPKNSLTPTNLNSTSIKTIBEKIELNSPK
LYNTDYSNLININFGNGRRIILKIPKNEFGLSLILLNNSIQNLKLDLNFHEL
NLTVLNDKYKINELLIDYRELKVKELHLYLLNPTIVNLYDSIYLHLVYSYGE
IIPNLKENLADFIKISTATFNKALKTISKQLEQPKLYSKPHSRETTIPLSISRE
YKSGMEIEPIFTYNDYQCFPHNKKIIOYLILKLSKTSSSSSNIEIEKEL
LNSINIEFLNFPNSMNELIKIADSNLSILSVEIKONLQIDMIFENCEKEFLKRSI
RNFKSQEPHFQFQNLNYSNLSSLPDRHYLPNNNNQDDDDTQIINSRMVLSY
PFIKSIRDCNTQMKLYSTKDDPFYFNITRLPAEQKQFFPSNLSDEEINPIQ
NLDNOPKFFPNLKVRIKSIIPORIKENELLKINFNDFIDEASHIFLAKSIEKLSI
NOIKLIGNGPPPIQDITSSSILLIRYKOPISPLSKCAKFSISDNCEKVTILOIFL
QMYENTKEDDDNDQDQAFHIIINHLKVLVSVKVQTIQVINKIYNOISSSSSS
SKHIIHSLHVVVYLSIKSPFLARYIMSRPFIYSITFDDDDGTTPYNCSDNPHQ
GYDIRKILNHLVFTPDFILGKTDHTTSFDLILEKLIDALSYPKMTDDESLLFTBEDY
LNE95"
18595..20304
/note="ORF_ID:dd_02710"
/codon_start=1
/product="PAXILLIN-LIKE PROTEIN"
/protein_id="AAM44368.1"
/db_xref="GI:21240657"
/translation="WATKGLNMDLDDLADLGRPKSSIKVTATVOTTATPSSGKNFD
NSDIONEIOSIIEELDQOQVOTISTPAKNHNTTTTASFVSQSPAPQPOQSQ
IDGLDDELMESENLSSTALKAVPTPEBIHANSPPSLKHTSNTSASSL
SRPNNSVSVTPQGVKVTSTITTTKQALPSKATLETTSGNNVYSQSPQSQOPY
KVTATNSQSDSDLDLGLSPSTTTTTPVPVORQOHHQHHQHHQHHQHHQHHQ
TQVTQINIGRTNPNNNNNNSPKVHGGDLDLNLNLTSQKDIDSTGPTSRG
TCGCRKPIFGETIQAMGKFYHPHFCHNCQDPLGTKNYQEBSELCHCEKCYQBLFC
ARCAHCDPISDRGITALGKWHVHVFCOCLKPEFGNPFERDGRPYCEADFYSTP
AVRCGCGSGIRGECINAGTQWHEHFVCOYCKSPGNGQPFEGGKPYCDVHVHQ
AGSCVCGCGKAVSGKVCVDAKDHPHFVCAFCNPLAGGSYIANGKPYCKGCHNK
LFA"
join (21782..22563,22679..22793,22936..27681)
/note="ORF_ID:dd_03308"
/codon_start=1
/product="putative protein contains similarity to
topoisomerase-related function protein"
/protein_id="AAM44369.1"
/db_xref="GI:21240658"
/translation="MDPKQVLSWVLSIGHELETLIDCNKDECDKLAMYSWKFO
DGYFDESVDIPGMPSPDSIDDDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
ANNIGDNNISNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
SSSSNNNNNSNGSTNNTSINSSSSSSSKLKSCHSGSIGNVVVCYSDIEMD
RRACSLKFKLEFLNDSIYHFENLQVFLPERKVEEKRIIVKLIRKSDTRFLG
TFTIESVSNQELVHLLTSLQSGFLQVCHLIDQSNWMSRPPKFWIFSFPFSLG
BFVAAPFESLIWAQFNNHIDGRKQSKSKQPNKPIPKLDSRSLNKEPMSKOWTI
DNREYAKAKGNLWHSINLNETPEKVDLDFIEKLGELIEDKSPSTSPSISSN
TTTTPTPTPTSTSTTTTTTTSSNNNNNNNNNNNNNNNNNNNNNNNNNNNN
RSTNNRNTNOSQQHVSVGGKLVGSSSSLESSENLSISSIPEETLVQNPYKLI
EFELCPSTERAGSTYDVTQGAQLQNFYSELSMDINSSENNRNRKIRGNNNS
GSLNSGNNPTVKQKKKQDQKQDKEKEKEKEKEKEKEKEKEKEKEKEKEKE
ANSLSNSGGGSKSLNDSSTQSDSTNRKTPANPTNTSIRKRRKKKSGRGN
NNNNNNNNOCGGGNSHGGHHHHHHSQNGKSSANSDEQESIQTRPRRSLP
MDTPPTIKGIFFLTTEDEQVMSLSESDIDHGKGTGPIKNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNSTIG"
```

```

FEATURES             source
Location/Qualifiers  1. 6457
                    /organism="Dictyostellium discoideum"
                    /mol_type="genomic DNA"
                    /strain="AX4"
                    /db_xref="taxon:44689"
                    /chromosome="2"
                    /map="209116-215573"
BASE COUNT          2444 a      616 C      632 G      2765 t
ORIGIN
Query Match          9.0%; Score 47.6; DB 3; Length 6457;
Best Local Similarity 64.5%; Pred. No. 2.8;
Matches 71; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy      61  TCATTATACCTTTATTATATCCCAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAAT 120
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5600 TCCTTTTAGTTTGGGATATTTTAAATTTTGTGATAAACCAAAATTTGATTTAAAAA 5541
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      121  AAATTTTCAAAACAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAG 170
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5540 AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA 5491
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 33
AC123513/C

```

DEFINITION	Dictyostelium discoideum chromosome 2 map 2779865-2840915 strain AX4, *** SEQUENCING IN PROGRESS ***.
ACCESSION	AC123513
VERSION	AC123513.1 GI:212420650
KEYWORDS	HTG; HTGS PHASE2.
SOURCE	Dictyostelium discoideum
ORGANISM	Dictyostelium discoideum
REFERENCE	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS	1 (bases 1 to 61052) Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P., Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.
TITLE	Sequence and Analysis of Chromosome 2 of Dictyostelium
JOURNAL	Unpublished
REMARK	The Dictyostelium Genome Sequencing Consortium
REFERENCE	2 (bases 1 to 61052)
AUTHORS	Baumgart,C.
TITLE	Direct Submission
JOURNAL	Submitted (29-MAY-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT	CDS predictions from GenElD may contain errors. Further Information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml) Funding Agency : Deutsche Forschungsgemeinschaft (DFG). * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.
FEATURES	Location/Qualifiers 1. .61052 /organism="Dictyostelium discoideum" /mol_type="genomic DNA" /strain="AX4" /db_xref="taxon:44689" /chromosome="2" /map="2779865-2840915" join(448..628,705..803,875..1134) /note="ORF_ID:dd_02708" /pseudo /codon_start=1 complement(join(1193..1583,1653..6117,6198..6330))
CDS	
CDS	

Matches	113;	Conservative	0;	Mismatches	109;	Indels	0;	Gaps	0;
Qy	26	ATTATATGCTGAATTTATATGGGTGAATTTTGCTTCCTTCTTTTACACTTTTATTTATTTCCCAA	85						
Db	37449	AATTTTAAGTGAATCATTTAAAAAACCTCACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	37390						
Qy	86	ATTTTCTTAAAGCAATAATTTCTTGGCTAATCAATAAATATCAGAAGAAAAAACAACG	145						
Db	37389	TTTTTTTTTTTACAAATATTAAAAATTAAATATACATAAATATTAAAAAATAAACAAACAA	37330						
Qy	146	AAAGCAACGCTTGAAAAAGGAAAGTAGCGCTATCGGCCTATCGGTATATTTTGGAGCTGTAAA	205						
Db	37329	AAAACAATAATTTATAAACAATAAATAAATAAACCAGTTATTTTTTTTATTTATTTATTTAT	37270						
Qy	206	TACTACRTGTTCTCTCTTAAGTCCCCTCTCTGTTTTCTTT	247						
Db	37269	TTTTTCTTTTGACITTTTTATATGAATCTTCAATGTAATAT	37228						
RESULT 36 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL REFERENCES AUTHORS TITLE JOURNAL COMMENT	AC004709 Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC004709 GI:4558585 HTG; HTGS PHASE1. Plasmodium falciparum (malaria parasite P. falciparum) Plasmodium falciparum Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W. Plasmodium falciparum 3D7 chromosome 12 Unpublished 2 (bases 1 to 196149) Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W. Direct Submission Submitted (21-WAY-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA On Apr 2, 1999 this sequence version replaced gi:4337173. NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	196149 bp DNA linear HTG 01-APR-1999							
FEATURES	source	Location/Qualifiers							
BASE COUNT	80057 a	17953 c	18800 g	77138 t	401 others				
ORIGIN									
Query Match	9.0%;	Score	47.6;	DB	2;	Length	196149;		
Best Local Similarity	50.9%;	Pred. No.	2.5;						
Matches	113;	Conservative	0;	Mismatches	109;	Indels	0;	Gaps	0;
Qy	26	ATTATATGCTGAATTTATATGGGTGAATTTTGCTTCCTTCTTTTACACTTTTATTTATTTCCCAA	85						
Db	62724	AATTTTAAGTGAATCATTTAAAAAACCTCACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	62783						
Qy	86	ATTTTCTTAAAGCAATAATTTCTTGGCTAATCAATAAATATCAGAAGAAAAAACAACG	145						

[illegible]

```
/locus_tag="PFL0935c"
complement(join(11613. .12941,13715. .19156))
/locus_tag="PFL0935c"
/notes="HMPfam hit to PF03011, Plasmodium falciparum
erythrocyte membrane protein (PFEMP), score 1.3e-56"
/codon_start=1
/product="erythrocyte membrane protein 1 (PFEMP1)"
/protein_id="AAN36275.1"
/db_xref="GI:23496720"
/translation="MVAAAGGGSSODAKHVLDEIGKIVYETVEKDAKLYKALKNGL
QKAKNGERASSNKTCTLYKEYEHFNGDANSRYPEKELSGKMGENRSDTLGGCT
NKIEGNKNKNGCAPYRHLCHLNHLESIDTTSITDILLAEVCYAAKPEGETLTFQ
HGOHQNTPGTASQCLCTVLARSADTIGVRGKDLFNGNDEKKRDELEENLTKIPA
KIHSGITSRGQNGKQRYKGDANKFYQLREDRWDRHRTWBEATCGAGESDKY
FROTACTGTTHKQCRDDKANVDPPTFYVPOYLRFWEAWEAFCKKNKKDKV
LNCNRGSDTDRYCSGNGDCTQITRAIKLVI GEHCTNCVWLRFKEWINOKLEF
LKQRNKQTEISGGGASGASGRSRTAATIKYEGYEKFEYEEPKKNGYGTVDLFL
GLNNKACKDINEEKEKIDFTKNVEDDKNINKEGTFYHSQYQCPGCGMKTKRNG
GWEEDTNCNKLKYKTSAPTEIKILSGEKHDDIKEKIEQFCKTQNGSDGSGGG
NSEKKELYDNWKYQIGELTKDQEGMDPDYDLVELTGGGLCILENKKKSDPKS
QNEPEQKTFNEFFYKFRFLNDSMYWRGVKNSCINKSKREKCEBCKEBCFCQK
WLGKKEWDKI KEHFKTDQFGNDALIGPEMKSPSVLKFVNLIDELFONIKSYGVNV
KELKGINKILDEKKEEERADGVVGASGDQNNNTIDKLLDHEDKATKCRNQ
EPTKASKEDLARSASDTPREDSPADLDDDEHDEDEDDDAESEESEEEKKE
EEHKEEENAEPEEKESGSSPEEVEEKAKEAPAGPACNIVOTLFSIDKFSDAC
TLKYGTAPTGWKCVTPSGDKAATSESGSGSDGAGDGVNGGALQRNKRDLATPS
AKSGDTTGGKDGATGSDGSI CVPPRRRLYIQKLHWA BKVGDATQPOVDTPSOD
KLRFDAIESAAVETFLMRYKKIKDKEKLEQORORENGELPGLSSGSDGSDNDPS
KLEKGIPOGFLRQMPYTLADYKIDLLYSGSDNTKSTYNDLIINGDKETQERKDIKG
ALSTYSNGETTTPPSPSVITPQWELHGPHIWNGMICALITYKENGIDKPEVDTA
VRAQLWDSNGTPNQEKYQYTNVKLEESAGAKINIPD7SGDNTPLNTPLTKNFVEI
PTPFRHLBWSGDSFCQRKRMKNVKNCRNI ERGGHEYCGDGHDCDTRDQHNML
ADPDCHDIOCRKRYKWKKEPEEHQNKTYKQKGLNGNSNDGNTKPKFETEN
RSPDQITSLHKCKNNGSGDPKNKIDFKEPLKTFGPLEYCKTCKPNKVCNGPSRR
SGNDQCTAVNGNEKIPSENGNSTTIDVHMDRGPYMEKSKOLENSENPLKT
SRLEGIHQKQWCKKQKNDIVCKLDQKQFQVDLQNYITTFIKVLLIWLDFIEGYIL
LKRRKIIETQKGEKTDDESNDCAVCQKVEKKTBEWQIKHEFKNRNOKGSDGN
DMKSSVRQLDPLIYMDLANGKINELKEFLSKYCECKVDNAGSEKDVVECLLQK
LETKAACKDPSSENGKTPCQBYTPDDEDDLLBEENTVKAPTCPKPKPEEQTE
ETCDKADEKVBETVAQENNI PPAGAPPADSPAGPATDSGKENVPEPPAPAP
RSPPIPLATSLAWSGIGFAATFYFLKKTKASVGNLQIOLIPKSDYDIPTLS
KNYIIPRSTYKGTIYMEGDADSKTYFMSDITDITSSSEISEBELDINDIYVPS
PKYKTLIEVLEPSENGNTASGNTPSDTQNDIQNDGIPSSKITDNEWNTLKDEFISN
MLQNKQKVPNDYSSGIDLPNTQNTLYFDNNQKPFITSIHDRNIYTGEEYSYNNM
VNNDNIPINRNDPSGIDLLINDSLANSKVDIYDELLKREKENELFGTEHPKHNTIYN
VAKPARDPTNNOINLFHKLDRHDMCEKMDTNKVDTILNOLKEWENDNSNGKNT
SGNITPSDIPSKLSDIPSTNKLMSDVSIQIHIDKPNQVDNIDYLDYPPKYTVDN
INPVDTHNPLVGNINPVDQNSNLTFPSNPNPAYDNITYIDHNEDLPKVKIEMSVK
NGEMAKEK"
repeat_region
12602. .12648
/rpt_type=tandem
repeat_region
12960. .13070
/rpt_type=tandem
repeat_region
12966. .13039
/rpt_type=tandem
repeat_region
13048. .13072
/rpt_type=tandem
repeat_region
13469. .13493
/rpt_type=tandem
repeat_region
13562. .13586
/rpt_type=tandem
repeat_region
16563. .16614
/rpt_type=tandem
repeat_region
17915. .17943
/rpt_type=tandem
repeat_region
19242. .19283
/rpt_type=tandem
repeat_region
19334. .19422
/rpt_type=tandem
repeat_region
20141. .20273
/rpt_type=tandem
repeat_region
20940. .21079
/rpt_type=tandem
```

```
Query Match          9.0%; Score 47.6; DB 3; Length 252650;
Best Local Similarity 50.9%; Pred. No. 2.4;
Matches 113; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 26 ATTATATGCTGATTTATGGGTGATTTTGGCTTCCTCTTTATACATTTTATTTATTTCCCAA 85
Db 90609 AATTTTATGTTGATGTATTCATTTAAAAACITTCACCTTTTTTTTTTTTTTTTTTTT 90668

QY 86 ATTTTCTTAAGCAATAATTTCTTTGCTAATCAATAAATATCAAAAGAAAAAAACATG 145
Db 90669 TTTTTTTTTTCAAAATATTAAATTAATATACATAATTAATAAATAAACAACAA 90728

QY 146 AAAGCAACGCTTTGAAAAAGGAAGCTTAGCCCTATCGGCTATATTTTGGAAAGTTGTA 205
Db 90729 AAACAATAATTTATAACAATAAATAAAGACCAGTTTATTTTATTTTATTTATTTAT 90788

QY 206 TACTACRGTGTTCTCTTAAAGTCCCACTCTCTGTTTCTTTT 247
Db 90789 TTTTCTTTTGTACTTTTTTATATGAATCTTCAATGTAAATAT 90830

RESULT 39
PFMAL1P2_1
WPCOMMENT
Sequence split into 4 fragments LOCUS PFMAL1P2 Accession AL031745
Fragment Name      Begin      End
PFMAL1P2_1         1          110000
PFMAL1P2_2         100001     210000
PFMAL1P2_3         200001     310000
PFMAL1P2_4         300001     384550
Continuation (2 of 4) of PFMAL1P2 from base 100001 (AL031745 Plasmodium falciparum DNA f.1)

Query Match          9.0%; Score 47.4; DB 3; Length 110000;
Best Local Similarity 52.8%; Pred. No. 2.8;
Matches 102; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 28 TTATATGCTGATTTATGGGTGATTTTGGCTTCCTCTTTATACATTTTATTTATTTCCCAAT 87
Db 36043 TAAATATATTATATAATATATATTTTTTTTTTTTTTTTACAAAATTTCCAGTAAAGCAATA 36102

QY 88 TTTTCTTAAGCAATAATTTCTTTGCTAATCAATAATATCAAAAGAAAAAAACATGAA 147
Db 36103 TATCCATTCGTATACAAATCTTTTCTTAATATACATATTCGAACAAAAAAAAGAA 36162

QY 148 AGCAACGCTTGAAAAAGGAAGTGTAGCCCTATCGGCTATATTTTGGAAAGTTGTA 207
Db 36163 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 36222

QY 208 CTACRTGTTCTCT 220
Db 36223 TTATCTATACTTTT 36235

RESULT 40
AX347077/c
LOCUS
DEFINITION
Sequence 2148 from Patent WO0200928.
ACCESSION
AX347077
VERSION
AX347077.1 GI:18494965
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.
REFERENCE
1
AUTHORS
Olek,A., Piepenbrock,C. and Berlin,K.
TITLE
Diagnosis of diseases associated with the immune system
JOURNAL
Patent: WO 0200928-A 2148 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
source
1. .113515
/organism="synthetic construct"
/mol_type="genomic DNA"
```


Qy 24 TGATTTATATGCTGATTTATTAAGGTGATTTTGCTCTCTTTATATCTTTATTTATTTATCC 83
Db 101048 TAAATTTTATCATATTTATTTATTTTATATATATTTATTTTCTGTTCAITTTTTTTTGA 101107
Qy 84 AAATTTTCTTAAAGCAATATTTCTTTGCTATCAATAATATCAAAAGAAAAAAAC 143
Db 101108 ATATTCATCTTATTTTATTTATATATTTTCTCATGATAGATTTTAAAAAGATAAAAAA 101167
Qy 144 TGAAGCAACGCTTGAAAAAGCAAGTTAGCCCTATCGGTATATTTTGGGAAGTTGTAA 203
Db 101168 ATAATACCTTTGTTAAATATATGTAACGATTTATAAATATGCAATATATGCAATATAAA 101227
Qy 204 AATACTACTGTCTCTCTCT 223
Db 101228 AATTTAATATACATCATGT 101247
RESULT 38
LOCUS AE014847 252650 bp DNA linear INV 11-FEB-2003
DEFINITION Plasmodium falciparum 3D7 chromosome 12, section 4 of 9 of the complete sequence.
ACCESSION AE014847
VERSION AE014847.1 GI:23496718
KEYWORDS
SOURCE
ORGANISM Plasmodium falciparum 3D7
Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 252650)
Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyes,S., Chan,M.S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguoli,S., Perlea,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
MEDLINE 22255705
PUBMED 12368864
REFERENCE 2 (bases 1 to 252650)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M., Nakao,B., Rowley,D., Tanaki,T., Wang,F. and Davis,R.W.
Direct Submission
TITLE Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
JOURNAL 3 (bases 1 to 252650)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M., Nakao,B., Rowley,D., Tanaki,T., Wang,F. and Davis,R.W.
Direct Submission
TITLE Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
JOURNAL Location/Qualifiers
FEATURES
source 1..252650
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"
repeat_region 90..134
/rpt_type=tandem
repeat_region 96..134
/rpt_type=tandem
repeat_region 586..669
/rpt_type=tandem
repeat_region 724..748
/rpt_type=tandem
repeat_region 1220..1250
/rpt_type=tandem
repeat_region 1331..1363

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
gene
mRNA
CDS
/rpt_type=tandem
1440..1476
/rpt_type=tandem
1758..1801
/rpt_type=tandem
1907..1971
/rpt_type=tandem
2150..2184
/rpt_type=tandem
2301..2344
/rpt_type=tandem
2484..2513
/rpt_type=tandem
<2608..>8601
/locus_tag="PFL0930w"
<2608..>8601
/locus_tag="PFL0930w"
2608..8601
/locus_tag="PFL0930w"
/note="HMPFam hit to PF00637, 7-fold repeat in Clathrin and VPS, score 2.1e-24"
/codon_start=1
/product="clathrin heavy chain, putative"
/protein_id="AAN36274.1"
/db_xref="GI:23496719"
/translation="MSQNPLSVCAVDNLINVDIQNESFRIGNVSVEGDKYICVKNV NENTQVVVNLHNKNSRKMKAESVLIHPNDPIALKGTIKNNTIPILOVFNLETKE KICSLNENYNYWKWINNDTIALVCEKNVYHNIDHTKKNKNDNNNNNNNNNN NNNNNFNTLTKVFERAKQIFDNNSQLIYSTDKMKWCLCGISTODQSGSDGYM QLYCEKLLHQITIEGFCFGFVDNLMKPLFCFVKKKNSTISRLHMLDIYNTK ETTPIYKVEINENLNDPFIYISLNTLQGVYIVTKGSYVVFDEGLSLIYKE KISDNIIFICDSNGENLIVAVNKKGIIFITINVPHLINHLNLSNFKDKIKNLK VKYGPQCDVISAAYKCIINDMPKASKILCLMNTKIPEHSNNIAESILIMPKNI DIRIIPERQOQVLNFKSPKNTSGQLSPILLYSVLLEYDKLNTYESIELKPVVLQ KKBYLEKIQDKLTCEELGVLDVLDURLALNIYLRCSHNKILUSTYCLLNMFN VLSYNNFQKINFQYIGVIFITVNYEFPQGNADKSAFVGDFSGSKNETMDFNDG AASISNQNNISINKLNGHINEVAIQIKFLCENNIISFDINKIIDYLSKKLCOEAT ILLDLENKAENKLOTKLPEFLYNNVOAETLFQMDIPTYDKNKIAYLCBEKGL YORALENYTINDIKRVITKSTCFQKGNNTTTTTSNMGDGHFDMNVSKGI SIEMKRYFTLSDSVCEELLDFPMKSKINMEVVISICVQYDKIGIKKLINFEEN KNYEGIFPYFYNILNPLNLIIVGNNDNNSLSSTYNNNDSDASILLTSDICSTP QYSYETANLKLIEDVHYIMFYIEACVKNINNIQELDRICKDKNAKPEQIKNFLCK LSPDRLIYVCDIHNYIEELAEYIKNSLLKYIEVYIVKPNNAHVIGVLLDLDA EDPLNLLNNIKNISNLIEIAEKNRLKLLPWLKESRANEGVENIELNALAKY IIDLKNPENLKNNFYDKKMGKYCEDLDPLAYTAYERSNGSCDEELIYITKNGL FLQARVLSRQSLNFKWLVLDENKRYKNVIDQVIGSTLIESNABEITVYKAFIE KKUSSELIELLEKIVLHNSFSDNNKLNLLILTAIKSDSKKVMYENIRLDFSGQPI AAVAYEYKLEBEAFVIYKFCNYTSAISVLDDKILYNNKQKNQNFHFSHSDTVPY SNNNNKNIIFGNKSDNENDEYLKQFTNNNSYREKYGSYNNNDISNNSNNIHLHPEF NNEKFSDDTLHNYEDIDSTESTCDLSNEDDLNRAIEFAQCNNDVWFILGKAOLK LNKIIDSISDFVKNPNPEAYKEVLEKCKENFVHLLTYLNTLREQNSLKDVLVDSL LYAVAKUKKNEMKFINTNSANLQIGRLYKEBEYEVAKILYSNIPNNKLTACY LKKEYALTEAAKKAASLTKWEVNIYKRYKQKYAHTAGLOIIMHADHLEIKI YEKKYIENLNLLENLNNERAHVIGYITGLIYAKYKPEKLMFIRNYTNKNTIKR LIDVCENEYLLKEAVLYISYDENLAVDTIIKHSPTAYTADTFQVHKVNTSDIIV KVIDFYIEEHPNLNLLKILENKIDNNRLVQTLKSNLPLKIOKYLEIDAQONTIAH NETNEIYQLNDQYISLRNSIDEDYDNFQNTNLNKLHKLPEWRITAAIYKKNKY KEALNLSKGEYKQKDAIEARVSKNNVTIEDLINFYIESKKNKEAFACLIYVDLLKP DYILEIVTSGQDAMLYFIQIISDYNTQIEHMKQIIEDEKEKKMKNSAPNDYSAM NNQFNYSLNNLSIMPPQNNFMSSNSFDKYDMFNNTHF"

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTA-250D13 is from a release of the human BAC library CITB-HS-A. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBeloBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The actual start of this clone is at base position 1 of RG250D13; actual end is at 176109 of RG250D13.

This clone contains STS sWSS2794 (NID:gl113588) and sWSS350 (NID:g484309).

FEATURES

Source

Location/Qualifiers

1..176109
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q31"
/clone="CTA-250D13"
/clone_lib="CITB-HS-A"

misc_feature

5770..5950
/notes="similar to EST T97069 (NID:g735693) ye50f03.r1"

misc_feature

5770..5949
/notes="similar to EST AW490098 (NID:g7060333)"

misc_feature

6407..6665
/notes="similar to EST W21257 (NID:gl298308) zb58b07.r1"

repeat_region

14504..14805
/rpt_family="L2"

repeat_region

14806..15107
/rpt_family="Alu"

repeat_region

15108..15194
/rpt_family="L2"

repeat_region

15360..15462
/rpt_family="MIR"

repeat_region

15633..15723
/rpt_family="L2"

repeat_region

15790..16074
/rpt_family="L2"

repeat_region

16787..16890
/rpt_family="MIR"

repeat_region

17143..17416
/rpt_family="MER1_type"

repeat_region

18425..18528
/rpt_family="MIR"

repeat_region

18853..18989
/rpt_family="Alu"

repeat_region

19024..19143
/rpt_family="Alu"

repeat_region

19338..19482
/rpt_family="MIR"

misc_feature

19664..20147
/notes="similar to EST AA034047 (NID:gi1505856) zi05e04.s1"

repeat_region

20436..20733

repeat_region
/rpt_family="Alu"
20898..21735
/rpt_family="L1"
repeat_region
21736..22031
/rpt_family="Alu"
repeat_region
22032..22079
/rpt_family="L1"
repeat_region
22690..23308
/rpt_family="L1"
repeat_region
25605..25729
/rpt_family="L2"
repeat_region
26006..26319
/rpt_family="Alu"
repeat_region
26377..26788
/rpt_family="L2"
repeat_region
28294..28390
/rpt_family="MIR"
repeat_region
30562..30741
/rpt_family="MIR"
repeat_region
30945..30984
/rpt_family="L2"
repeat_region
31078..31146
/rpt_family="L2"
repeat_region
32776..32830
/rpt_family="L2"
repeat_region
33004..34302
/rpt_family="L1"
repeat_region
34389..34576
/rpt_family="L2"
repeat_region
35613..35812
/rpt_family="MIR"
repeat_region
35980..36553
/rpt_family="L1"
repeat_region
36887..37180
/rpt_family="Alu"
repeat_region
37807..37949
/rpt_family="Alu"
repeat_region
38095..38581
/rpt_family="MaLR"
repeat_region
38876..39185
/rpt_family="Alu"
repeat_region
39402..39485
/rpt_family="L2"
repeat_region
39826..39926
/rpt_family="MIR"
repeat_region
41270..41572
/rpt_family="Alu"
repeat_region
41637..41939
/rpt_family="Alu"
repeat_region
42310..42463
/rpt_family="Alu"
repeat_region
42583..42791
/rpt_family="MER1_type"
repeat_region
43575..43952
/rpt_family="CR1"
misc_feature
44284..44598
/note="similar to EST AW946860 (NID:g8124633)"
repeat_region
44407..44534
/rpt_family="Alu"
repeat_region
44582..44701
/rpt_family="CR1"
repeat_region
44716..44926
/rpt_family="CR1"
repeat_region
47256..47583
/rpt_family="L2"
repeat_region
47635..48161
/rpt_family="L2"
repeat_region
49690..49883
/rpt_family="L2"
repeat_region
50136..50338
/rpt_family="L2"

Query Match

9.0%; Score 47.4; DB 9; Length 176109;

```

/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT      35172 a      1174 c      22520 g      54649 t
ORIGIN

Query Match
Best Local Similarity 54.4%; Pred. No. 2.8;
Matches 93; Conservative 1; Mismatches 77; Indels 0; Gaps 0;

QY 55 CTTCTCTTTTATCTTTTATTTTCCCAATTTTCTTAAGCAAAATTTCTTTGCTA 114
|||||
Db 41662 CTTCTCTTTTATCTTTTATTTTCTTAAGCAAAATTTCTTTGCTA 114
|||||

QY 115 ATCAATAAATATCAAAAGAAAAAACTGAAGCAAGCTTTGAAAAAGGAAAGTTAG 174
|||||
Db 41602 ATTATATATTATCACTATAAATCTATAAATCTATAAATCTATAAATCTATAAATCTCA 41543

QY 175 CCTATCGGTATATTTGGAGTTGTAAATCTACRTGTCTCTCTTAA 225
|||||
Db 41542 TTCCCAACATTTATTTTATATATCAATTTATCAAAATTTAACTTAA 41492

RESULT 41
AC004153
LOCUS      156060 bp      DNA      linear      HTG 12-AUG-2000
DEFINITION Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
ACCESSION  PROGRESS ***, 2 unordered pieces.
VERSION     AC004153
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM    Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1 (bases 1 to 156060)
AUTHORS     Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tanaki,T.,
            Kurdi,O.B., Conway,A.B. and Davis,R.W.
TITLE       Plasmodium falciparum 3D7 chromosome 12
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 156060)
AUTHORS     Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
TITLE       Direct Submission
JOURNAL     Submitted (18-FEB-1998) Stanford DNA Sequencing and Technology
            Center, Stanford University, 855 California Avenue, Palo Alto, CA
            94304, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            * 1 102169: contig of 102169 bp in length
            * 102170 102369: gap of unknown length
            * 102370 156060: contig of 53691 bp in length.
FEATURES
source      Location/Qualifiers
            1..156060
            /organism="Plasmodium falciparum"
            /mol_type="genomic DNA"
            /db_xref="taxon:5933"
            /chromosome="12"
            /clone="PPVAC812"
            /clone="3D7"

BASE COUNT      62615 a      14889 c      15137 g      63219 t      200 others
ORIGIN

Query Match
Best Local Similarity 52.9%; Pred. No. 2.7;
Matches 99; Conservative 1; Mismatches 87; Indels 0; Gaps 0;

QY 39 TTTATGGTGATTTTGGTCTCTTATATCTTTATTTATTTATTTTCCCAATTTTCTTAAGC 98
|||||

```

```

Db 69975 TATATCGCTACTGAAGTATGCCTATTATTATCTATTATTATTTTACAAATATATTTTATAT 70034
QY 99 AAATATTTCTTTGCTAAATCAATAATATTATCAAAAGAAAAAACTGAAAGCAACGCTTG 158
|||||
Db 70035 GGATATAATATTTTGGCGCTGAACAATAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 70094
|||||
QY 159 AAAAAGGAAAGTTAGCCCTATCGGTATATTTTGGAGTTGTAAAAATCTACTACRTGTCT 218
|||||
Db 70095 ATTATAAATATATTTTATTTATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTCT 70154
|||||
QY 219 CTTCTAA 225
Db 70155 ATTATAA 70161

RESULT 42
AC003992/c
LOCUS      176109 bp      DNA      linear      PRI 04-JUN-2002
DEFINITION Homo sapiens BAC clone CTA-250D13 from 7q31, complete sequence.
ACCESSION  AC003992
VERSION     AC003992.2 GI:21322221
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 176109)
AUTHORS     Sulston,J.E. and Waterston,R.
TITLE       Toward a complete human genome sequence
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063792
PUBMED      9847074
REFERENCE   2 (bases 1 to 176109)
AUTHORS     Minx,P., Hinds,K., Sutterer,C., Becker,M. and Ozersky,P.
TITLE       The sequence of Homo sapiens BAC clone CTA-250D13
JOURNAL     Unpublished (2001)
REFERENCE   3 (bases 1 to 176109)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (14-JAN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4 (bases 1 to 176109)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (03-FEB-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5 (bases 1 to 176109)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (04-FEB-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6 (bases 1 to 176109)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (04-JUN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jun 4, 2002 this sequence version replaced gi:2772533.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_RG250D13

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate

```


QY	39	TTTATGGGTGATTTCTCTTCTTATCTTTATTTATTTATTTCCCAAAATTTTCTTAAGC	98
Db	88878	TATATGCGTACTAGTATGCGTATTTATCTATTTATGTTTACAAATATATTTTATAT	88937
QY	99	AAATATTTCTTGTCTAATCAATAAATTTATCAAAAGAAAAAATGAAAGCAACGCTTG	158
Db	88938	GGATATAAATTTTGTAGCGGTACAAATAGAAAAGAAAAAAGAAAAAAGAAAAA	88997
QY	159	AAAAAGGAAGCTTACCGCTATCGGTATATTTTGGAAAGTTGTAAATACTACRGTCTT	218
Db	88998	ATTAATAAATATATTTATTTATTTGTTTAAATGTTATTTTATTAATAATATTTATGTT	89057
QY	219	CTTCTAA	225
Db	89058	ATTATAA	89064
RESULT 45			
PFMAL3P7			
LOCUS			
DEFINITION			
ACCESSION			
AL034559	AL008974	AL008975	AL008981
AL010143	AL010146	AL010154	AL010157
AL010187	AL010189	AL010207	AL010208
AL139179	AL844502	Z98556	Z98557
AL034559	4	GI:8052273	
HTG:	40S Ribosomal protein S3A; acyl transferase;		
KEYWORDS	acylaminoacyl-peptidase; ATP-dependent RNA Helicase; cyclophilin; elongation factor; F49C12.11-like protein; HeSB-like domain protein; histone H2A variant; kinesin-related protein; N-acetylglucosamine-1-phosphate transferase; protein kinase; R-CG7 repeat; R-PA3 repeat; rep11; rep20; rifin; RNA-binding protein; stevor; T-complex protein 1 epsilon subunit; telomere; var.		
SOURCE	Plasmodium falciparum 3D7		
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
REFERENCE	1 (bases 1 to 253305)		
AUTHORS	Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Gentles, C., Gwilliam, R., Hamlin, N., Harris, D., Holtwell, T., Hornsby, T., Horrocks, P., Jagers, K., Jassal, B., Kyes, S., McLean, J., Moul, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and Barrell, B.G.		
TITLE	The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum		
JOURNAL	Nature 400 (6744), 532-538 (1999)		
MEDLINE	99376085		
PUBMED	10448855		
REFERENCE	2		
AUTHORS	Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Akin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holtroyd, S., Horrocks, P., Humphray, S., Jagers, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moul, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinovitch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrell, B.G.		
TITLE	Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13		
JOURNAL	Nature 419 (6906), 527-531 (2002)		
MEDLINE	22255708		
PUBMED	12368867		
REFERENCE	3 (bases 1 to 253305)		

AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

Lawson, D., Bowman, S. and Barrell, B.
Direct Submission
Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
On or before May 14, 2001 this sequence version replaced
gi:2982540, gi:2982541, gi:2982544, gi:2894453, gi:2894464,
gi:2982551, gi:2982556, gi:2982558, gi:2982563, gi:2982564,
gi:2982566, gi:2894491, gi:2982569, gi:2695974, gi:2894502,
gi:2673766, gi:2894496, gi:2982577, gi:2894588, gi:2894589,
gi:2982538, gi:2982539, gi:2894380, gi:4725992.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.

Location/Qualifiers
1..253305
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="3"
/clone="MAL3P7"
458..4060
/gene="MAL3P7.1"
/note="MAL3P7.1"
/note="Similarity: to kinesin-related protein.
D.melanogaster kinesin-like protein 67A (TR:P91945) BLAST
Score: 664, sum P(2) = 1.9e-67; 28% identity in 707 aa
overlap; Pfam: match to PF00225 kinesin (6.1e-86), Kinesin
motor domain Score 311.54 E-value 6.1e-86"
/codon_start=1
/product="kinesin, putative"
/protein_id="CAB39023.1"
/db_xref="GI:4493364"
/db_xref="GOA:O97277"
/db_xref="SPTREMBL:O97277"
/translation="MRKTNLNDPFPFLNNYYNNKEENPIIICSDSEKERTKIYSTNH
HTNITLYNDHTLKICNTENNKAHLHSNKSLOKNILMPSYMKQKTHRETIN
VYPNTGSPSTVENITGEHPINGQYDALKNMNLNYDHOHNTMNNISNKKLPVL
NNNNNQDVIENNVEYPTSKNIYDSIYIPIQINIKITINSEENNNNNINNDNNH
NTYINNNNNNNNPTSVTSKNNNNNNLINTLNAYSNVAVARIKPIGSEENIVS
IFNKYVLIKEKEKEXLLSQKQSTYFDSFVDVNAQTEVEFFQTAKLPIPHVFK
GINCTVFAYGATSGKTYMLDDKQNGIVQLSLELFTIINEKCRNIKVLMSFLEV
NETTRDLGKEKNKTLVEQDAEYKVNLSLIEIVNVEQALINIEGVKRMKSP
RANKVSRSHAIQIYVYNEILDNMNTISYKAKLCFVDLAGSASASNSKGRKE
GSYINQSLALANCINSLASNRISKVVKYRDSKTLHKNSLEGNCIVMTANLP
STFSQESNNTUKYAFRANIKCATVQINDKESDIEKILKKNENLQKEYDTLLGY
TNLKEFFFTIINQLYKKQSCYKLIENISDNMSMELKQDITDYDLVKMKSDSYR
KKVDSLKOLYQEEKQLNNLFDTPLEKLNLYVINSKVDNNKSLLEEMIFPKHNNK
VNENFLVNEKVVQKNNLVNGVVDENVVQKNNLVNGVVDENVVVDENVVVDENV
VNDENVVDENVVDENVVDENVVDENVVDENVVDENVVDENVVDENVVDENV
ADGKSESHNNKNDIEDNDKTIKIDHNNNNSSDNDDEYOSANSVPESDVIKKEKK
KKIPINMETKGTQWGTQKDPHKPTDYDINVLGINKEDVSNKSNNDYNTNKNIEKNY
EKGGYNPPHNNLTQMSILYIINNENNVSPHSPMKKNVAKMLKGLNLTANFIL
DDDTTKNNNSKISDKNNNNILNENKIKDKSKCKCKNNNNNNNNNNNNNNNN
NNNNNNSSSSGKVDGINLNSNTNERLHTFGVYSLNLDNETKIKIKKMEKN
DHLTSTIDTISKHARDLLKNRKLNFQENIKHEKDEVSLEYVKKKIKKKOM"
join(5457..5577,5689..5787,5923..6077,6189..6743)
/gene="MAL3P7.2"
/note="synonym: PFC0865w"
join(5457..5577,5689..5787,5923..6077,6189..6743)
/gene="MAL3P7.2"
/note="Similarity eg to C.elegans RNA-binding protein
(TR:Q18318) BLAST Score: 378, sum P(1) = 3.8e-35; 37%
identity in 297 aa overlap; pfam: aplicing conserved in P.
knowlesi and P. yoelii; Pfam: match to PF00076 frm, RNA
recognition motif. (a.k.a. RRM, RBD, or RNP domain) Score
58.01, E-value 1.1e-20"
/codon_start=1
/product="RNA-binding protein, putative"
/protein_id="CAB39067.1"
/db_xref="GI:4494008"

THIS PAGE BLANK (USPTO)

```
/db_xref="GOA:O97318"
/db_xref="SPTREMBL:O97318"
/translation="MGLFDKIRNIEKLNKAEALKNIGNDSSWHDOYRDDSSYIYIGNLD
NRLTGGDIVVPSQVDPVNDVNDNETGSKGYCFLSYADQDSTILAVNENGKYL
LERPLVDHILNVPKPKYKLDADKNEYKPTGAEGQGLGVNVVSESEIKLSKVPDKIK
KNSNEKKLLDDELDWALNPEKSIKDIISPIGDEKSRNHEGMEDEDEDD
DEDDDDSDSDIKYIKRHKRKSUTTKYIKRHKRKSODHRDKRHRNHRHREK
EKDKSKHRHRHSHYDYSRSRSTSSSTR"
join(7801..7928,8084..8234,8392..8583)
/gene="MAL3P7.3"
/note="synonym: PFC0870w"
join(7801..7928,8084..8234,8392..8583)
/gene="MAL3P7.3"
/note="Similarity to P. falciparum elongation factor
1-beta, (AAF27524) (276 aa), fasta scores: opt: 366, E():
1.1e-15, 53.2% identity in 141 aa overlap, revised:
shortened exon 2, splicing confirmed in P. knowlesi; Pfam:
match to PF00736 rF1BD, EF-1 guanine nucleotide exchange
domain Score 61.10, E-value 7.9e-22"
/codon_start=1
/product="elongation factor 1 (EF-1); putative"
/protein_id="CAB39068.2"
/db_xref="GI:8052274"
/db_xref="GOA:O97319"
/db_xref="SPTREMBL:O97319"
/translation="MANTYDELYPLSYIILQNEGNTSKIDQANTKKPKKEVINKS
SLIIDIPYGENTDLDVLKVNITMEGLTWGAHKKTPEAFGLFKLQVSCVIVDDL
VNTDLIEIENLGNLOKQKQMDDEYDEDDDEIGGLVQSAEIIISFKL"
join(10596..10737,11044..11146,11373..20416,20578..20690)
/gene="MAL3P7.4"
/note="synonym: PFC0875w"
join(10596..10737,11044..11146,11373..20416,20578..20690)
/gene="MAL3P7.4"
/note="Revised: added 3' exon, possibly spliced at 5' end,
revised: added 2x 5' exons by similarity with P. knowlesi;
signal anchor predicted by SignalP 2.0 HMM (Signal peptide
probability 0.001, signal anchor probability 0.997) with
cleavage site probability 0.000 between residues 36 and 37
Pfam match to entry PF00005 ABC_tran, ABC transporter,
score 72.50, E-value 9e-18"
/codon_start=1
/product="transporter, putative"
/protein_id="CAB39024.3"
/db_xref="GI:23477031"
/db_xref="GOA:O97278"
/db_xref="SPTREMBL:O97278"
/translation="MKIKLYGLLKKIYKRDFFLYFFFLIIPFLIISPHLFIRNI
SDKYSLEYNTSSPNDKINLNDTIQVYFLSNELCSEINDEIYVNHICLTPGNELM
KDFUSYANNDLINIFKLYNSEDCLNKYAIELKSEDFKVNLSLELNDLINTHEND
KNSLLYNLHTNIINLKSEFNIDLYKDYITENEFEINIKLNIPHLKVVYKTEFYK
NFIKDIKNLKNENHYQVFNDDRKKLFFNFVKNLVETKYSCGLSVENVVYKQK
NVKPSYPLSPQSESGKGLYKNAPLYNKYENKOYGNPFLYYPVYVLESDDKINE
KNERNEKNERNEKNERNEKNGKNGKNDILSSTNNNVVKNQIQSYHNNNSIN
DDQKKNNNNYIYINPDGLITNVKIRKVDYALLNSNEKGFNDININLKQNF
NFNTVDDLSFNINFEWYFFVILEYQFNSFLYNADILKKNLLTPRYDEINEL
NKEBETMEYKTKDEKYQVNNYINENVGDEKGYKTELKLNDFIYKPMKMK
INAPDFEKNIFRVVFLVCLFVNIQFIDNKERKINENFLCCLKINKYVYVSWL
LFYFILLPFYNIIFTLVVYVYVYKLMNYFILLVFFLMNSLFTIICWQFSNNS
INYATFLLPFLSSFLIIHSGASNLTFVFLIIPHSSFCLSLDFILIKNIKID
YKQIFPENISLMLHIIICSIVSFLLICILNIYIYKRNKNNMLNTRKGSKITN
DKRKSQQTVEYGINYDSEKYEYFQHKDQNSDFMLGSDDEDDCYELDTLHKQV
NSOMCLRLEKIRNDRRESNGHPDIPMLPLRNNNNINNNNNNNNNNNNNNN
NOHRNNOCYLIQVNNPYGKHALKQVSLTRSNRIFFVLGNGSGKSTLINTKM
ISKDSGEINFPKSYKNNKDIYMDVKNRNGLCGRICSGYIKFVWNRIPKNCN
DTNYDNNNNNNNDHNNHNSNGINSFNNNNVHKNKNDLEISYCSQNVIL
YENITFYETIKIFLIYTNKNDKRLTKIMNDLQVYLNQDLKMLDEVKKKI
SIFICFLVKRDIYILDEPFDALDKTKLKFDFDKIKNNIIFICTHIEANNFAN
DIAVKSQIIFKGSQKSLIEYKFTLANIQFNGLPNEQNTYLSQEKIMSALNKG
VIGSPRVANNYVDGVNVRISSESSINDKNDNNNDNNNNNNNNNNNNNNNN
ILVTNINEFTNRDI IKNNIIKFTGTRENKNCIFPENENHYCTYKINELESKKL
LCVLNKPKNILTYQLKTDIYTYIYITLNEKKKLLAKIQDDKDIKYLIEIDPLFLP
FQNFYFNLNKLMLKNNHMQPISYNFAYLRDILVRGNENKNNIINSTNDDNDPYDN
MYNNHNVGKYKNNNNNSFYNMRDDTDELTDIENISSKKKKKFKIEKFSYNTYRN
```

```
KNNDENDEYGYDNDYDDNIMSNSTSNIMKKKNIKKQNIKFTCINFFTTYIKPTLLKL

Query Match
Best Local Similarity 9.0%; Score 47.4; DB 3; Length 253305;
Matches 96; Conservative 1; Mismatches 82; Indels 0; Gaps 0;

QY 38 ATTATGGGTGATTTTGCTTCCTTCTTTATACCTTTTATTTATCCCAAATTTTCTTAAG 97
DB 82760 ATATATATATATATATATATTTTATTTATTTATTTATTTATTTATTTATTTT 82819

QY 98 CAATATTTCTTCTGCTAATCAATAATATCAAAAGAAAAAACTGAAAAGCAAGCTT 157
DB 82820 TATTTTCTTCTTATGAAAAAAAATGAAGCCCAAGGAAAAACATAATGAAATATA 82879

QY 158 GAAAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAATACCTACTGTGT 216
DB 82880 AAATATAAATATAAAAAATAAAAAATTCATATTTATATAATAATAATAAGAAAT 82938

Search completed: February 14, 2004, 16:11:49
Job time : 2730 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 14:16:25 ; Search time 2621 Seconds
(without alignments)
4905.406 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggatgctgtagtctgt.....tcttgagaggctcttgag 529

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estcov:*

6: em_estpl:*

7: em_estro:*

8: em_est1:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gsa_hum:*

18: em_gsa_inv:*

19: em_gsa_pln:*

20: em_gsa_vrt:*

21: em_gsa_fun:*

22: em_gsa_nam:*

23: em_gsa_mus:*

24: em_gsa_pro:*

25: em_gsa_rod:*

26: em_gsa_phg:*

27: em_gsa_vrl:*

28: gb_gsa1:*

29: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	481.6	91.0	515	28	B56032 CIT-HSP-200
C 2	434.2	82.1	564	28	AQ790276 HS_2240_B
C 3	368.6	69.7	395	28	AQ803995 HS_2005_A
4	60	11.3	1091	13	AX424950 BX424950

5	57	10.8	936	13	BX404443
c	55.6	10.5	770	9	AL514205
6	55	10.4	252	13	BX425020
7	54.6	10.3	894	28	BH137168
8	54.2	10.2	427	13	BX403499
c	53.8	10.2	1201	13	BX446296
c	53.4	10.1	506	9	AL514791
c	53.4	10.1	1162	9	AL513719
c	53	10.0	234	13	BX425129
c	52.8	10.0	1104	9	AL513809
c	52.6	9.9	252	13	BX425020
c	52.4	9.9	1201	9	AL514421
c	52.4	9.9	1203	29	CNS015WU
c	52.2	9.9	699	13	BX424825
c	52.2	9.9	1201	13	BX446296
c	52	9.8	934	13	BX425394
c	51.8	9.8	923	13	BX425397
c	51.6	9.8	964	14	CB756565
c	51.4	9.7	512	13	BX424931
c	51.2	9.7	589	29	CNS030I2
c	51.2	9.7	738	9	AL513907
c	51.2	9.7	838	10	BP264841
c	51.2	9.7	856	12	BM15165
c	50.8	9.6	927	13	BX371218
c	50.6	9.6	234	13	BX425129
c	50.6	9.6	381	13	BX425197
c	50.6	9.6	894	9	AL543011
c	50.4	9.5	332	13	BX446492
c	50.4	9.5	381	13	BX425197
c	50.4	9.5	940	13	BX404361
c	50.4	9.5	1201	9	AL513975
c	50.2	9.5	375	13	BX404101
c	50.2	9.5	661	9	AL514015
c	50.2	9.5	1024	13	BX349610
c	50.2	9.5	1126	13	BX446388
c	50.2	9.5	1162	9	AL513719
c	50.2	9.5	1189	9	AL513823
c	50.2	9.5	1201	9	AL513597
c	50	9.5	928	29	CNS00DKY
c	49.8	9.4	294	13	BX446415
c	49.8	9.4	314	13	BX446274
c	49.8	9.4	417	13	BX465145
c	49.8	9.4	484	14	CA722141
c	49.8	9.4	505	9	AL513553
c	49.8	9.4	535	13	BX446937
c	49.8	9.4	633	13	BX403519
c	49.8	9.4	738	9	AL513907
c	49.8	9.4	865	9	AL515307
c	49.8	9.4	1063	13	BX414736
c	49.6	9.4	358	12	BM514261
c	49.6	9.4	513	13	BQ264913
c	49.6	9.4	663	12	BJ441961
c	49.6	9.4	665	12	BJ423198
c	49.6	9.4	1009	9	AL514865
c	49.6	9.4	1179	13	BX425126
c	49.4	9.3	356	13	BX446758
c	49.4	9.3	886	9	AL514691
c	49.4	9.3	909	9	AL514871
c	49.4	9.3	1123	9	AL513999
c	49.2	9.3	292	12	B1742580
c	49.2	9.3	422	9	AW813554
c	49.2	9.3	1126	13	BX446388
c	49.2	9.3	1193	9	AL513783
c	49.2	9.3	1201	9	AL514529
c	49	9.3	197	14	CA746319
c	49	9.3	281	12	BM521967
c	49	9.3	297	13	BX425620
c	49	9.3	409	13	BX425137
c	49	9.3	425	28	AA550654
c	49	9.3	602	13	BX404150
c	49	9.3	959	13	BX403428
c	49	9.3	1056	13	BX415058
c	49	9.3	1065	9	AL515373

```

78 49 9.3 1104 9 AL513809
79 49 9.3 1189 12 BM547682
80 49 9.3 1201 9 AL551891
81 49 9.3 1998 11 BC025253
82 48.8 9.2 557 13 BQ390979
83 48.8 9.2 866 9 AL515321
84 48.8 9.2 924 9 AL515235
85 48.8 9.2 1056 13 BX415058
86 48.8 9.2 1136 9 AL514047
87 48.8 9.2 1201 29 CNS016CN
88 48.6 9.2 279 13 BX425800
89 48.6 9.2 289 13 BX425107
90 48.6 9.2 320 13 BX446411
91 48.6 9.2 329 12 BM063269
92 48.6 9.2 366 13 BX446397
93 48.6 9.2 399 9 AL514627
94 48.6 9.2 561 13 BQ395163
95 48.6 9.2 613 28 AQ22590
96 48.6 9.2 888 28 AZ549422
97 48.6 9.2 898 13 BX432881
98 48.6 9.2 899 13 BX453223
99 48.6 9.2 909 9 AL514871
100 48.6 9.2 994 13 BX414650

```

ALIGNMENTS

```

RESULT 1
B56032/c
LOCUS
DEFINITION
  B56032 CIT-HSP-2005M18.TR CIT-HSP Homo sapiens genomic clone 2005M18,
  genomic survey sequence.
ACCESSION
  B56032.1 GI:2610366
VERSION
  B56032
KEYWORDS
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 515)
  Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden
  ,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M.
  and Venter,J.C.
  Use of a random BAC End Sequence Database for Sequence-Ready Map
  Building
JOURNAL
  Unpublished
COMMENT
  Other_GSSs: CIT-HSP-2005M18.TF
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com) . BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
  Seq primer: M13 Reverse
  Class: BAC ends.

```

```

FEATURES
  source
    1..515
      Location/Qualifiers
        1..515
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
          /clone="2005M18"
          /sex="Male"
          /cell_type="Sperm"
          /clone_lib="CIT-HSP"
          /notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
          HindIII"
BASE COUNT
  152 a 102 c 112 g 149 t

```

```

ORIGIN
  Query Match 91.0%; Score 481.6; DB 28; Length 515;
  Best Local Similarity 98.6%; Pred.No.3e-56; Mismatches 1; Indels 1; Gaps 1;
  Matches 495; Conservative 1;
  1 CATGGTGATGCTGTAGTCTGGCTGATTATATATGCTGATTTATGGGTGATTTTGGCTTCCT 60
  502 CATGGTGATGCTGTAGTCTGGCTGATTATATGCTGATTTATGGGTGATTTTGGCTTCAT 443
  61 TCTTTATATCTTTTATTTATCCAAATTTTCTTAAGCAAAATATTTCTTTGCTAATCAAT 120
  442 TCTTTATATCTTTTATTTATCCAAATTTTATTAAGCAAAATATTTCTTTGCTAATCAAT 383
  121 AAATTATCAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 179
  382 AAATTATCAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 323
  180 TCGGGTATATTTTGGAAAGTTGTAAATATCTACTGTCTTCTTCTTAAGTCCCACTCTCTCTG 239
  322 TCGGGTATATTTTGGAAAGTTGTAAATATCTACTGTCTTCTTCTTAAGTCCCACTCTCTCTG 263
  240 TTTTCTTTGAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 299
  262 TTTTCTTTGAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 203
  300 CGGCTGCTTGCCTGCTCCATCACAGAAAGGTGCAGCAGAGCCCTTTTCTTCCAGGGTTG 359
  202 CGGCTGCTTGCATGCTCCATCACAGAAAGGTGCAGCAGAGCCATTTTCTTACAGGGTTG 143
  360 TGGGCACCTTGAACATATCCTGGCTTGTAGCTTAAGAAATGCCCCAAGTGAAGCAGCAGTC 419
  142 TGGGCACCTTGAACATATCCTGGCTTGTAGCTTAAGAAATGCCCCAAGTGAAGCAGCAGTC 83
  420 AAGCTGTTGGGGGACTTCCAGAAATAAACCAACAGCTGTATTTACTGTCCTTTGACATCC 479
  82 AAGCTGTTGGGGGACTTCCAGAAATAAACCAACAGCTGTATTTACTGTCCTTTGACATCC 23
  480 TCTCTGCTCAGGGCCCAAGTCAC 501
  22 TCTCTGCTCAGGGCCCAAGTCAC 1
  RESULT 2
  AQ790276/c
  LOCUS
  DEFINITION
    HS_2240_B1_F01_MR CIT Approved Human Genomic Sperm Library D Homo
    sapiens genomic clone Plate=2240 Col=1 Row=L, genomic survey
    sequence.
  ACCESSION
    AQ790276
  VERSION
    AQ790276.1 GI:5697900
  KEYWORDS
    GSS.
  SOURCE
    Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 564)
    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
    Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
    Hood,L.
    Sequence-tagged connectors: A sequence approach to mapping and
    scanning the human genome
    Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  JOURNAL
    99380589
  MEDLINE
    10449764
  PUBMED
    Contact: Mahairas GG, Wallace JC, Hood L
    High Throughput Sequencing Center
    University of Washington
    401 Queen Anne Avenue North, Seattle, WA 98109, USA
    Tel: (206) 616-3618
    Fax: (206) 616-3887
    Email: jwallace@u.washington.edu
    Clones may be purchased from Research Genetics (info@resgen.com).

```

BAC end Web Server: <http://www.htsc.washington.edu>
 Plate: 2240 row: L column: 1
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 564.

FEATURES

E-COLL DHI08"				
BASE COUNT	150 a	110 c	128 g	145 t
ORIGIN				
				31 others

Query Match	82.1%	Score 434.2;	DB 28;	Length 564;
Best Local Similarity	95.7%	Pred. No. 7.8e-50;		
Matches 455; Conservative	1;	Mismatches 17;	Indels 3;	Gaps 2;

45 GGTGATTTTGGCTTCCTTCTTTATACTTTTATTTTATTTCCCAATTTTCTTAAGCAAATAT 104

Db

564 GGCGATTATGCTTCCTCTTTATACCTTTTATTGATGCCAAATTTCTTTAGCAGATAT 505

Qy 105 TTCTTTTGGCTAATCAATAAATTATCAAAAGAAAAAAACTGAAAGCAACGCTTGAAAAAA 164

Db 504 TTCTTTTGCTAATCAATATAATTATCAAAAGAATTATAAAACCTGAGAGCAACGCTTGAAAAA 415

Qy 165 GGAAAGTTAG-CCCTATCGGGGTATATTTTGGAAAGTTGTAATACTACTACTGTCTCTTCT 223

Db 444 G G A N A G T T A G C C C C T A T C G G G T A T A T T T T G G A A G T T G T A A A A A T A C T A C G T G T T C T C T T C T 385

Qy 224 AAGTCCCACTCCTCTGTGTTTCTTTTGAGCAGGAAAGAGAACAGATCACCTCTCTTCCAT 283

Db
384 AAGTCCCACTCCTCTGTTTCCTTTGAGCAGGAAGAAGACAGTACCCCTCTCTTCCAT

284 GACAAACCCATGACGGCTGTTGCTCTCCATCACAGGAAGTGCAGCAGACCCCT 343

[illegible][illegible]

404 CAACTGACGGCACTCAAGCTGTTCGGGCACTTCAGAGATAAACCAGCCAGCTATTA 463

Db 204 CAAGTGAGCCACACTCAAGCTGTTGGGACCTTCCAGATAAACCACCAAGCTGTATTA 145

QY 464 CTGGTCTTTGACATCCTCTCTGTCTCAGG--CCAAGTCACGTGTCCTCTGTCTTGGAGAGGG 521

D_b 144 CTGGTCTTTGACATCCTTAAGAACAGAGAAACAAGTCNCTGTCCCTGTCTTGGAGAGG 85

Qy 522 CTTCTG 527

Db 84 CTGTG 79

RESULT 3

AQ803995/c
LOCUS
AQ803995 395 bp DNA linear GSS 09-AUG-1999
DEFINITION
HS 2005_A2_G09 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone plate=2005 Col=18 Row=M, genomic survey sequence.

```

VERSION      BX424950.1  GI:30770416
KEYWORDS
SOURCE       EST.
ORGANISM     Homo sapiens (human)

REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      1 (bases 1 to 1091)
              Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 6403.r For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CL0BA005ZA01FP1&cluster=6403.r. Contact :
              Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CL0BA005ZA01FP1.
              Location/Qualifiers
              1..1091
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CL0BA005ZA01"
               /tissue_type="PLACENTA"
               /note="Vector: pcMVSPORT 6; 1st strand cDNA was primed
               with a NotI-oligo(dT) primer. Five prime end enriched,
               double-strand cDNA was digested with Not I and cloned into
               the Not I and EcoRV sites of the pcMVSPORT 6 vector.
               Library was not normalized."
BASE COUNT   252 a 96 c 41 g 303 t 399 others
ORIGIN
Query Match 11.3%; Score 60; DB 13; Length 1091;
Best Local Similarity 30.5%; Pred. No. 32;
Matches 71; Conservative 74; Mismatches 88; Indels 0; Gaps 0;

QY 50 TTTTGCTCTCTCTTTATATCTTATTTATTCCTCAAAATTTTCTTAAGCAAAATATTTCTT 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 TTTTITTTTTTDAAAAAATTTTTTTTTTWRRRGGGTTTTTDARRAAAAAANWTTT 634

QY 110 TGTATCAATCAATATATCAAGAAAAAATCTGAAGCAAGCTTGCAAAAAAGGAAA 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 TTTWAAAAAATTTTTTTAAAAAARRRRAAAAAAARRRRRRRRAAAAAA 694

QY 170 GTTAGCCCTATCGGTATATTTTGAAGTTGTAATACTACTGTCTCTCTAAGTCC 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 695 AAATTTWRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRTTTTKTKTKGTAWDDNW 754

QY 230 CACTCTCTCTGTTTCTTGAGCAGGAAAGCAAGCAGTCACCTCTCTTCCA 282
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 755 WTTTTTTTTTTTTTTTDRRRRRRRRRRRRRRTTTTWHMMWMMWMMWMM 807

RESULT 5
BX404443
LOCUS        BX404443 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION   CL0BB029ZG02 3-PRIME, mRNA sequence.
ACCESSION    BX404443
VERSION      BX404443.1  GI:30635181
KEYWORDS
SOURCE       EST.
ORGANISM     Homo sapiens (human)
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 936)
              Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
              Library was constructed by Life Technologies, a division of

Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6537.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BB029ZG02FP1&cluster=6537.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BB029ZG02FP1.
Location/Qualifiers
1..936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB029ZG02"
/tissue_type="NEUROBLASTOMA"
/note="Vector: pcMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pcMVSPORT 6 vector.
Library was not normalized."
BASE COUNT   270 a 33 c 113 g 326 t 194 others
ORIGIN
Query Match 10.8%; Score 57; DB 13; Length 936;
Best Local Similarity 46.8%; Pred. No. 88;
Matches 111; Conservative 18; Mismatches 108; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTTATCGGTGATTTTCTCTCTCTCTTTATATCTTTATTTATCCCAA 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 TTTTITTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 93

QY 87 TTTTCTTAAGCAATATTTCTTCTCAATCAATTAATATCAAGAAAAAATCACTCA 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 TTTTITTTTTTTTTTTTTTTTTTTTTTTTWWAAAAAATTTTAAAMWMAAAATTTTA 153

QY 147 AAGCAACGCTTGAAAAAGAAAGTAGCCCTATCGGTATATTTTGGAAAGTTGTAAAT 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 AAGGAATWATTAATAAAAWAAAAAATTCSSSTTAAAAAATAAAATTTTWWAAA 213

QY 207 ACTACRTGTCTCTCTAAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 263
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 214 WKTITTTTGTGGGGTTTWWAAAAAATTTTTTTTTTTTTTWWAAAAAATTTGGA 270

RESULT 6
AL514205/c
LOCUS        AL514205 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION   CL0BB006ZG12 3-PRIME, mRNA sequence.
ACCESSION    AL514205
VERSION      AL514205.2  GI:30464090
KEYWORDS
SOURCE       EST.
ORGANISM     Homo sapiens (human)
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 770)
              Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
              Full-length cDNA libraries and normalization
              Unpublished
              On Feb 13, 2001 this sequence version replaced gi:12777699.
              Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of

```

Invitrogen. This sequence belongs to sequence cluster 6437.r For more information about this cluster, see <http://www.genoscope.cns.fr/> cgi-bin/cluster.cgi?seq=CL0BB0062G12P1&cluster=6437.r. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CL0BB0062G12FP1. Location/Qualifiers

FEATURES

source

1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB0062G12"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
Location/Qualifiers
216 a 35 c 193 g 181 t 145 others

Query Match 10.5%; Score 55.6; DB 9; Length 770;
Best Local Similarity 52.7%; Pred. No. 1.5e+02;
Matches 77; Conservative 17; Mismatches 52; Indels 0; Gaps 0;

Qy 24 TGATTATATGCTGATTTATGGTGATTTTGGTCTCTTATATCTTTATTTATTTATCC 83

Db 277 KTK 218

Qy 84 AAATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAA 143

Db 217 TTTTCTTAAADWAATTTTTTYYAAAAA 158

Qy 144 TGAAGCAAGCTTGAAAAAGGAAA 169

Db 157 AAAAAA 132

RESULT 7
BX425020 252 bp mRNA linear EST 15-MAY-2003
DEFINITION BX425020 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA007ZD06
3-PRIME, mRNA sequence.

ACCESSION BX425020

VERSION BX425020

KEYWORDS BX425020.1 GI:30772428

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 252)

Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 8225.r For

more information about this cluster, see

<http://www.genoscope.cns.fr/>

cgi-bin/cluster.cgi?seq=CL0BA007ZD06FP1&cluster=8225.r. Contact :

Feng Liang Email : liang@lifetech.com URL :

<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CL0BA007ZD06FP1.

Location/Qualifiers

1..252

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

FEATURES

source

/clone="CL0BA007ZD06"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
Location/Qualifiers
84 a 20 c 24 g 89 t 35 others

FEATURES

source

Query Match 10.4%; Score 55; DB 13; Length 252;
Best Local Similarity 47.9%; Pred. No. 3.2e+02;
Matches 79; Conservative 23; Mismatches 63; Indels 0; Gaps 0;

Qy 5 GTGATGCTGCTAGTCGTGCTGATTTATATGCTGATTTATGGTGATTTGCTCTCTTCTT 64

Db 68 GGGATACCTAGAGCGCGCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 127

Qy 65 TATATCTTTTATTTATTCCTCAAAATTTTCTTAAGCAATATTTCTTTGCTAATCAATAAT 124

Db 128 TT 187

Qy 125 TATCAAAAGAAAAAAGCAAGCAAGCTTGAAAAAGGAAA 169

Db 188 WAAAAA 232

RESULT 8

BH137168

LOCUS BH137168

DEFINITION ENTOKI4TR Entamoeba histolytica Sheared DNA linear GSS 07-AUG-2001

genomic, genomic survey sequence.

BH137168

ACCESSION BH137168.1 GI:15096229

VERSION GSS.

KEYWORDS Entamoeba histolytica

SOURCE Entamoeba histolytica

ORGANISM Eukaryota; Entamoebidae; Entamoeba.

1 (bases 1 to 894)

Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.

Determination of clone end sequences from Entamoeba histolytica

HMI:IMSS sheared DNA library (2001)

Unpublished

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Reverse

Class: shotgun

High quality sequence start: 24

High quality sequence stop: 522.

Location/Qualifiers

1..894

/organism="Entamoeba histolytica"

/mol_type="genomic DNA"

/strain="HMI:IMSS"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: PHOS1; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,


```
QY 326 AAGTGTGACGACGAGCCCTTTCTTCCAGG 355
      : : : : : : : : : : : : : : : :
Db 854 WMCACACGBCGCCCCYVYATATCSGS 825

RESULT 18
BX424825/c
LOCUS
DEFINITION
BX424825 Homo sapiens PLACENTA Homo sapiens cDNA clone
XLOCBA0012E02 3-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 699)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : XLOCBA001ZE02FP1.
Location/Qualifiers
1..699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="XLOCBA001ZE02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 172 a 88 c 92 g 113 t 234 others
ORIGIN
1..699
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="XLOCBA001ZE02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
Query Match 9.9%; Score 52.2; DB 13; Length 699;
Best Local Similarity 53.1%; Pred. No. 4.6e+02;
Matches 76; Conservative 14; Mismatches 53; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTATGGGTGATTTGCTTCTTATATCTTTTATTTATTTCCCAA 86
      ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 283 TTTTITTTKGGTTTTTTTKKTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTHCCCM 224
      ||| | | | | | | | | | | | | | | | | | | | | | | | |

QY 87 TTTTCTTAAGCAATATTTCTTTGCTTAATCAATAAATTATCAAAAGAAAAAACTGA 146
      ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 223 TTTTITTKGAATKGAAKTTTTTTTTTTTAYAAAAATTTTTTTWWRRRAAAAAA 164
      ||| | | | | | | | | | | | | | | | | | | | | | | |

QY 147 AAGCAACGCTTGAATAAGGAA 169
      ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 163 AAAAAAATAAAAAAATAAAAAA 141
      ||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 19
BX446296
LOCUS
DEFINITION
BX446296 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
XLOCBB001ZA03 3-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1201 bp mRNA linear EST 15-MAY-2003
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6437.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=XLOCBB012G07FP1&cluster=6437.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : XLOCBB001ZA03FP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="XLOCBB001ZA03"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 663 a 37 c 191 g 116 t 194 others
ORIGIN
1..1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="XLOCBB001ZA03"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
Query Match 9.9%; Score 52.2; DB 13; Length 1201;
Best Local Similarity 31.5%; Pred. No. 3.5e+02;
Matches 45; Conservative 55; Mismatches 43; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTATGGGTGATTTGCTTCTTATATCTTTTATTTATTTCCCAA 86
      ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 148 TTTTITTTTWWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 207
      ||| | | | | | | | | | | | | | | | | | | | | | | |

QY 87 TTTTCTTAAGCAATATTTCTTTGCTTAATCAATAAATTATCAAAAGAAAAAACTGA 146
      ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 208 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 267
      ||| | | | | | | | | | | | | | | | | | | | | | | |

QY 147 AAGCAACGCTTGAATAAGGAA 169
      ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 268 AAATATATATATAAAAAAATAAAAAA 290
      ||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 20
BX425394/c
LOCUS
DEFINITION
BX425394 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
XLOCBB012ZG07 3-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 934)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6437.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=XLOCBB012G07FP1&cluster=6437.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

Faraday Avenue Genoscope sequence ID : CLOBB012ZG07FP1.

FEATURES

source
1. 934
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB012ZG07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 228 a 184 c 192 g 142 t 188 others
ORIGIN

Query Match 9.8%; Score 52; DB 13; Length 934;
Best Local Similarity 50.6%; Pred. No. 4.2e+02;
Matches 80; Conservative 16; Mismatches 62; Indels 0; Gaps 0;

Qy 12 TGTAGTCGCTGCTGATTATATGCTGATTATGCTGATTATGCTGCTTCTTATACCT 71
Db 281 TTWADTTTTTKKKKKKTTTTTTTTTTTATGCTGCTTCTTATACCT 222
Qy 72 TTATTTATCCCAATTTTCTTAAGCAAAATATTTCTTGTCTTAATAAATATCAAA 131
Db 221 TTT 162
Qy 132 AGAAAAAAACTGAACGACGCTTGAAGGAAA 169
Db 161 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 124

RESULT 21

LOCUS BX425397/c
DEFINITION BX425397 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB012ZH03 3-PRIME, mRNA sequence.

ACCESSION BX425397

VERSION BX425397.1 GI:30774457

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 923)

Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqreg@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4504.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CLOBB012ZH03FP1cluster=4504.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/InvitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CLOBB012ZH03FP1.

FEATURES

source
1. 923
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB012ZH03"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 369 a 140 c 113 g 196 t 105 others
ORIGIN

Query Match 9.8%; Score 51.8; DB 13; Length 923;
Best Local Similarity 53.6%; Pred. No. 4.5e+02;
Matches 98; Conservative 4; Mismatches 81; Indels 0; Gaps 0;

Qy 26 ATTATATGCTGATTATGCTGATTATGCTGCTTCTTATACCTTTATTTATCCCAA 85
Db 231 ATTATGTTTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 172
Qy 86 ATTTTCTTAAAGCAAAATATTTCTTGTCTTAATAAATATCAAAAGAAAAAACTG 145
Db 171 WTTTTCCTTTTAAAGTTTTTTTTTTTTTTTTTAAAGAAAAAAAGAAAAAA 112
Qy 146 AAAGCAACGCTTGAAGAAAGAAAGTAGCCCTATCGGCTATATTTGGAAGTTGTA 205
Db 111 AAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGTWTCCCTCGAGGGCCCA 52
Qy 206 TAC 208
Db 51 TAC 49

RESULT 22

LOCUS CB756565
DEFINITION AGENCOURT 12978927 NICHD_XGC_Tad1 Xenopus laevis cDNA clone
IMAGE:6877304 3', mRNA sequence.

ACCESSION CB756565

VERSION CB756565.1 GI:29836072

KEYWORDS EST.

SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 964)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Sequenced with vector primer

Tissue Procurement: Drs. Donald Brown and Liqian Cai

cDNA Library Preparation: CLONTECH

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM3115 row: 1 column: 07

High quality sequence stop: 158.

Location/Qualifiers

1. 964

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:6877304"

/dev_stages="metamorphosis stage 53"

/clone_lib="NICHD XGC_Tad1"

/note="Organ: Developing Tadpole; Vector: pDNR-LIB;

Site 1: Sfi; Site 2: Sfi; 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence:

5'-CACGCCCATATGCC-3' and 3' adaptor sequence:

5'-ATTCTAGAGCGGCGGCGCATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G or T). Average insert size 1.6 kb

(range 0.9-3.0 kb). 15/15 colonies contained inserts by

PCR. This library was enriched for full-length clones and

was constructed by Clontech Laboratories (Palo Alto, CA)."

[illegible][illegible]


```

DEFINITION AL543011 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YL22
ACCESSION 5-PRIME, mRNA sequence.
VERSION AL543011
SOURCE EST.
KEYWORDS AL543011.2 GI:30548727
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 894)
COMMENT Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12875489.

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9876.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE013DF11Q1&cluster=9876.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE013DF11Q1.

FEATURES
source
1..894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YL22"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 234 a 158 c 203 g 237 t 62 others
ORIGIN
Query Match 9.6%; Score 50.6; DB 9; Length 894;
Best Local Similarity 41.5%; Pred. No. 6.7e+02;
Matches 81; Conservative 33; Mismatches 81; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTATGAGGATTTGCTTCCTTATACCTTTTATTTATTTATTTCCAAA 86
Db 691 TTTTCTTTTCTTAAAGAAAGTGTAGCCCTATCGGCTATATTTTGGAGTTGTAAAT 750
QY 87 TTTTCTTAAAGAAATTTCTTGTCTAATCAATTAATTAAGAAAGAAAGAAAGTGA 146
Db 751 TTTTATTTTCTTTTATTTTATTTTAAKAAAAAATAAAWAAWAAATAAAAAA 810
QY 147 AAGCAACGCTTGAAAGAAAGTGTAGCCCTATCGGCTATATTTTGGAGTTGTAAAT 206
Db 811 AAAAAAARARARARARARARARARARARARARARARARARARARARARARAR 870
QY 207 ACTAGTGTCTCTT 221
Db 871 ATTAAKAWTTTTTT 885

RESULT 32
BX446492/c
LOCUS
DEFINITION BX446492 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA0092C08
3-PRIME, mRNA sequence.
ACCESSION BX446492
VERSION BX446492.1 GI:31033732
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 332)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6609.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA0092C08P1&cluster=6609.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA0092C08P1.

FEATURES
source
1..332
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA0092C08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 143 a 25 c 24 g 122 t 18 others
ORIGIN
Query Match 9.5%; Score 50.4; DB 13; Length 332;
Best Local Similarity 53.2%; Pred. No. 1.2e+03;
Matches 84; Conservative 9; Mismatches 65; Indels 0; Gaps 0;

QY 12 TGTACTCGTGGCTGATTATATGCTGATTATGCGGTGATTTGCTTCCTTTATACCT 71
Db 292 TTTKTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 233
QY 72 TTATTTATTTCCAAATTTTCTTAAGCAAAATTTTCTTGTCTAATCAATAAATATCAAA 131
Db 232 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 173
QY 132 AGAAAAAAGCAAGCAAGCTTGAAAAAGGAAA 169
Db 172 AAAAAAAGCAAGCAAGCTTGAAAAAGGAAA 135

RESULT 33
BX425197
LOCUS
DEFINITION BX425197 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB003ZH07 3-PRIME, mRNA sequence.
ACCESSION BX425197
VERSION BX425197.1 GI:30769941
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 381)
COMMENT Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6437.r For

```


[illegible]

RESULT 43
CNS00DKY
LOCUS
DEFINITION
 CNS00DKY 928 bp DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR27A24 of RPCI-38 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION AL071865
KEYWORDS AL071865.1 GI:4948170
SOURCE
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 928)
AUTHORS Genoscope.

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : secre@genoscope.cns.fr)
- Web : www.genoscope.cng.fr

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pierer de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES
    source
        Location/Qualifiers
            1..328
                organism="Drosophila melanogaster"
                mol_type="genomic DNA"
                db_xref="taxon:7227"
                clone="BACR27A24"
                clone_lib="RPC1-98"
                note="end : 77"
            262 a 70 c 84 g 321 t 191 others

```

Query Match 9.5%; Score 50; DB 29; Length 928;
Best Local Similarity 31.3%; Pred. No. 7.9e+02;
Matches 62; Conservative 59; Mismatches 77; Indels 0; Gaps 0

Qy	50	TTTTGCTTCCTTCATTATACCTTTTATTTATCCCAATTTTCTTAAGCAATAATTCTT	109
Db	711	TTTTTTTTTTTWTATWAAWATAWAATAATWWWWWATWTDWGKNNNNAWWWWWWWWW	770

Qy 110 TGCTAATCAATAATATCAAAAGAAAAAACTGAAAGCAACGTTGAAAAAGGAAA 169

Db 771 AWWWWAAAAWWWWWWWWWWAAAAAAWAAAAAAWDDDDDDDDDDWWKAAKKKK 830

Qy 170 GTTAGCCCTATCGGCTATATTTTGGAGTTGTAAATACTACRTGTCTCTTCTAAGTCC 229

Db 831 KKKKKKKKKKKKBGKKKTATTTTGGAGTTGTAAATACTACRTGTCTCTTCTAAGTCC 890

Qy	230	CACTCCCTCTGTTTTCTTT	247
Db	891	TTTTTTTTTTTTTTTTTTTT	908

RESULT 44	
BX446415/c	
LOCUS	DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

COMMENT

FEATURES
SOURCE

BASE COUNT

BASE COUNT	Query Match	Best Loc	Matches
ORIGIN			

Matches

Qy

Db

Qy Db Qy

Qy
Db
PESINT 45

RESULT 45
BX446274/c
LOCUS
DEFINITION
ACCESSION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BX446415 294 bp mRNA linear EST 22-MAY-2003
 BX446415 Homo sapiens PLACENTA Homo sapiens CDNA clone CLOBA0052H05
 3-PRIME, mRNA sequence.
 BX446415
 BX446415.1 GI:31029724
 EST.

201. Homo sapiens (human)
202. Homo sapiens
203. Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
204. Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 294)

Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Séquençage

genoscope+ Centre national de séquençage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5229.r For
more information about this cluster, see

more information about this cluster, see
[http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CL0BA005ZH05P1&cluster=5229.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CL0BA005ZH05P1&cluster=5229.r). Contact :
 Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CL0BA005ZH05P1

```

1. .294
   Location/Qualifiers
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"

```

```

/cd_xrefs" taxon="9506"
/clone="CLOBA005ZH05"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-soligo(4T) primer five prime end enriched

```

100 a 18 c 19 g 130 t 27 others

ch 9.4%; Score 49.8; DB 13; Length 294;
1 Similarity 52.4%; Pred. NO. 1.5e+03;
75; Conservative 13; Mismatches 55; Indels 0; Gaps 0;

27 TTTTATATGCTGATTATGCGTGATTTGGCTTCCTTCCTTATACTTTTATTATTCCAAA 86
||| | : || : ||| : ||| : ||| : ||| :
72 TTTTTTTTTKKKTTTTKKKTATTTTTTTTTTKTKTTTTTTTTTTTTTTTTTTTTTTTTVVVY 213

87 TTTTTTCTTAAAGCAATATATTTCTGAATCCTAATTAATTCATGACGAAAAAACAATCA 146

[illegible][illegible]

BX4456274 314 bp mRNA linear EST 15-MAY-2003
 BX4456274 Homo sapiens PLACENTA Homo sapiens CDNA clone
 XCLOBA0012B11 3-PRIME, mRNA sequence.
 BX4456274
 BX4456274 1 CT-20776405

BX446274.1 GI:30776405
 EST.
 Homo sapiens (human)
 Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 314)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Contact: Genoscope
Genoscope - Centre National de Sequencage

GENOSCOPE - CENTRE NATIONAL DE
BP 191 91006 EVRY cedex - France

BF 191 91006 EVRI Cedex - France
Email: segref@genoscope.cns.fr Web : www.genoscope.cns.fr

Email: seqlife@genoscope.cns.fr, web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of

Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1769.f. For

more information about this cluster see

more information at
<http://www.denoscon>

http://www.genoscope.cns.fr/
cqi-bin/cluster_cqi?seq=YCLOB00017B11EP1.fc]net or = 1769 f Contact :

cgi-bin/ciuster.cgi?seq=xCLUBA001ZBIFFP&club

Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen demonstration 1600

<http://fulllength.invitrogen.com/> Invitrogen Corporation
 6500 Central Expressway, Fremont, CA 94558-5080
 Tel: 415 352 3000 Fax: 415 352 3001
 Email: fulllength@invitrogen.com

Faraday Avenue Genoscope sequ

Location
1 314

1.314

```
/organism="Homo sapiens"
```

/mol_type="mRNA"

/db_xref="taxon:9606"

```
/clone="XCL0BA001ZB11"
```

/tissue_type="PLACENTA"

```
/clone_lib="Homo sapiens PLACENTA"
```

/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and clone

the Not I and EcoRV sites of

Library was not normalized."

BASE CC

9.48: Query Match Score 49.8: DB 13: Length 314:

Query Match 5.4%, score 43.0, DB 13
Best Local Similarity 58.7%: Pred. No. 1.5e+03:

Best Local Similarity 58.7%; Fied: NO: 1.3e+03;
 Matches 84: Conservative 1: Mismatches 58: Indels 0: Gaps 0:

OV 27 TTTATATGCTGATTATGGGTGATTGTGCTTCTTCTTTATACCTTTATTTATTCGCCAA 86

QY Z/ TTTATATGCTGATTTAAGGGTGATTTGGCTTCTCTTATATACCTTTATTTATTCCTGAAA 89

250
 309
 250

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

D_b

190

Search completed: February 14, 2004, 16:55:33
Job time : 2632 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 13:45:13 ; Search time 325 Seconds
(without alignments)

5995.821 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggtgatgtgtagtctg.....ctctggagagggcttctgag 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubnpa/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubnpa/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubnpa/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubnpa/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubnpa/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubnpa/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubnpa/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubnpa/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubnpa/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubnpa/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubnpa/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubnpa/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubnpa/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubnpa/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubnpa/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubnpa/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubnpa/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubnpa/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.6	99.9	529	13	US-10-027-632-1
2	528.6	99.9	529	14	US-10-027-632-1
3	527.6	99.7	535	13	US-10-027-632-54345
4	527.6	99.7	535	13	US-10-027-632-300106
5	527.6	99.7	535	14	US-10-027-632-54345
6	527.6	99.7	535	14	US-10-027-632-300106
7	51.4	9.7	321	13	US-09-814-353-17876
8	51	9.6	284	13	US-09-814-353-17553
9	49.6	9.4	424	10	US-09-960-352-11218
10	48.2	9.1	316	13	US-09-814-353-16036
11	47.4	9.0	113515	13	US-10-311-455-2148
12	47.2	8.9	597	13	US-09-814-353-4872
13	47.2	8.9	597	13	US-09-814-353-11169
14	46.6	8.8	312	10	US-09-960-352-8414
15	46.6	8.8	383	13	US-09-814-353-18006

46.4	8.8	480	12	US-10-131-827-8192	Sequence 8192, Ap
46.2	8.7	312	10	US-09-960-352-8414	Sequence 8414, Ap
46	8.7	3673778	13	US-10-312-841-2	Sequence 2, Appli
45.8	8.7	463	13	US-09-814-353-4849	Sequence 4849, Ap
45.8	8.7	463	13	US-09-814-353-11146	Sequence 11146, A
45.8	8.7	3673778	13	US-10-312-841-1	Sequence 1, Appli
45.6	8.6	406	13	US-09-814-353-17782	Sequence 17782, A
45	8.5	359	13	US-09-814-353-17359	Sequence 17359, A
44.8	8.5	321	13	US-09-814-353-17876	Sequence 17876, A
44.8	8.5	723	13	US-09-814-353-4726	Sequence 4726, Ap
44.8	8.5	723	13	US-09-814-353-11024	Sequence 11024, A
44.4	8.4	228	13	US-09-814-353-17588	Sequence 17588, A
44.4	8.4	425	10	US-09-834-975-451	Sequence 451, App
44.4	8.4	664	9	US-09-739-254-66	Sequence 66, Appli
44.4	8.4	664	13	US-10-055-988-66	Sequence 66, Appli
44.4	8.4	664	15	US-10-054-988-66	Sequence 66, Appli
44.4	8.4	6078	13	US-10-311-455-1110	Sequence 1110, Ap
44.2	8.4	3673778	13	US-10-312-841-1	Sequence 1, Appli
44	8.3	281	13	US-09-814-353-17383	Sequence 17383, A
44	8.3	600	15	US-10-198-846-8434	Sequence 8434, Ap
43.8	8.3	236	10	US-09-960-352-12183	Sequence 12183, A
43.8	8.3	281	13	US-09-814-353-17383	Sequence 17383, A
43.8	8.3	610	13	US-09-814-353-5380	Sequence 5380, Ap
43.8	8.3	610	13	US-09-814-353-11667	Sequence 11667, A
43.6	8.2	9884	15	US-10-172-086-5	Sequence 5, Appli
43.6	8.2	6590	13	US-10-311-455-1449	Sequence 1449, Ap
43.4	8.2	424	10	US-09-960-352-11218	Sequence 11218, A
43.2	8.2	380	10	US-09-960-352-9335	Sequence 9335, Ap
43.2	8.2	397	13	US-09-814-353-17585	Sequence 17585, Ap
43.2	8.2	430	15	US-10-078-090-87	Sequence 87, Appli
43	8.1	255	13	US-09-930-213-715	Sequence 715, Ap
43	8.1	512	13	US-09-814-353-5072	Sequence 5072, A
43	8.1	512	13	US-09-814-353-11364	Sequence 11364, A
43	8.1	180557	14	US-10-003-806-6	Sequence 6, Appli
43	8.1	180557	14	US-10-003-806-9	Sequence 9, Appli
43	8.1	368004	10	US-09-949-654-3	Sequence 3, Appli
43	8.1	3673778	13	US-10-312-841-2	Sequence 2, Appli
42.8	8.1	358	15	US-10-198-846-9603	Sequence 9603, Ap
42.8	8.1	383	13	US-09-814-353-18006	Sequence 18006, A
42.6	8.1	515	13	US-09-814-353-6142	Sequence 6142, Ap
42.6	8.1	515	13	US-09-814-353-12421	Sequence 12421, A
42.4	8.0	308	10	US-09-960-352-7670	Sequence 7670, Ap
42.4	8.0	384	13	US-09-814-353-2534	Sequence 2534, Ap
42.4	8.0	384	13	US-09-814-353-8870	Sequence 8870, Ap
42.4	8.0	524	13	US-10-125-968-1258	Sequence 1258, Ap
42.4	8.0	1081	13	US-10-027-632-118342	Sequence 118342, A
42.4	8.0	1081	13	US-10-027-632-118343	Sequence 118343, A
42.4	8.0	1081	14	US-10-027-632-118343	Sequence 118343, A
42.4	8.0	1081	14	US-10-027-632-118343	Sequence 118343, A
42.4	8.0	6731	13	US-10-311-455-1033	Sequence 1033, Ap
42.2	8.0	277	10	US-09-960-352-12673	Sequence 12673, A
42.2	8.0	277	10	US-09-960-352-12673	Sequence 12673, A
42.2	8.0	302	13	US-09-814-353-5335	Sequence 5335, Ap
42.2	8.0	302	13	US-09-814-353-11622	Sequence 11622, A
42.2	8.0	380	10	US-09-960-352-9335	Sequence 9335, Ap
42.2	8.0	390	13	US-09-814-353-17808	Sequence 17808, A
42.2	8.0	418	10	US-09-960-352-4845	Sequence 4845, Ap
42.2	8.0	435	13	US-09-814-353-5103	Sequence 5103, Ap
42.2	8.0	435	13	US-09-814-353-11395	Sequence 11395, A
42.2	8.0	476	13	US-09-814-353-4700	Sequence 4700, Ap
42.2	8.0	476	13	US-09-814-353-10999	Sequence 10999, A
42.2	8.0	550	13	US-09-814-353-5098	Sequence 5098, Ap
42.2	8.0	550	13	US-09-814-353-11390	Sequence 11390, A
42.2	8.0	1066	9	US-09-770-445-141	Sequence 141, App
42.2	8.0	5359	13	US-10-311-455-1580	Sequence 1580, Ap
42.2	8.0	5359	13	US-10-240-453-146	Sequence 146, App
42.2	8.0	5359	13	US-10-240-453-257	Sequence 257, App
42.2	8.0	5359	13	US-10-240-453-4844	Sequence 4844, Ap
42.2	8.0	5359	13	US-09-814-353-11141	Sequence 11141, A
42.2	8.0	5359	13	US-09-814-353-4872	Sequence 4872, Ap
42.2	8.0	5359	13	US-09-814-353-11169	Sequence 11169, A
42.2	8.0	5359	13	US-10-311-455-2303	Sequence 2303, Ap

QY 301 GGCTGCTTGCCTGCTCCATCATCAGGAAGGTGACAGAGCCCTTTTCTTCCAGGGTTGT 360
Db 301 GGCTGCTTGCCTGCTCCATCATCAGGAAGGTGACAGAGCCCTTTTCTTCCAGGGTTGT 360
QY 361 GGGCACCTTGAACATATCTCGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCA 420
Db 361 GGGCACCTTGAACATATCTCGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCA 420
QY 421 AGCTGTTCCGGGACTTCCAGAAATAACCAACACGCTGTATTACTGCTTTTGACATCCT 480
Db 421 AGCTGTTCCGGGACTTCCAGAAATAACCAACACGCTGTATTACTGCTTTTGACATCCT 480
QY 481 CTCTGGTCAGGCCCAAGTCACTGTCCTCTGTCTTGAGAGGGGCTTCTGAG 529
Db 481 CTCTGGTCAGGCCCAAGTCACTGTCCTCTGTCTTGAGAGGGGCTTCTGAG 529

RESULT 3

US-10-027-632-54345
; Sequence 54345, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54345
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54345

Query Match 99.7%; Score 527.6; DB 13; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-119;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGTGATGCTGTAGTCGTGCTGATTTATATGCTGATTTATGGTGATTTGCTTCCTT 61
Db 1 ATGGTGATGCTGTAGTCGTGCTGATTTATATGCTGATTTATGGTGATTTGCTTCCTT 60
QY 62 CTTTATACCTTTTATTTATCCAAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATA 121
Db 61 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATA 120
QY 122 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAGTTAGCCCTATC 181
Db 121 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAGTTAGCCCTATC 180
QY 182 GGGTATATTTTGGAAAGTTGTAAAACTACTACGTGTTCTCTTCTTAAGTCCCACTCTCTGTT 241
Db 181 GGGTATATTTTGGAAAGTTGTAAAACTACTACGTGTTCTCTTCTTAAGTCCCACTCTCTGTT 240
QY 242 TTCTTTGAGCGAAGAAAGACGATCACCCCTCTCTTCCATGACAAACCAATGACCG 301
|||||

Db 241 TTCTTTGAGCGAAGAAAGACGATCACCCCTCTCTTCCATGACAAACCAATGACCG 300
QY 302 GCTGCTTGCCTGCTCCATCATCAGGAAGGTGACAGAGCCCTTTTCTTCCAGGGTTGTG 361
Db 301 GCTGCTTGCCTGCTCCATCATCAGGAAGGTGACAGAGCCCTTTTCTTCCAGGGTTGTG 360
QY 362 GGCACCTTGAACATATCTCGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCA 421
Db 361 GGCACCTTGAACATATCTCGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCA 420
QY 422 GCTGTTCCGGGACTTCCAGAAATAACCAACACGCTGTATTACTGCTTTTGACATCCT 481
Db 421 GCTGTTCCGGGACTTCCAGAAATAACCAACACGCTGTATTACTGCTTTTGACATCCT 480
QY 482 TCTGGTCAGGCCCAAGTCACTGTCCTCTGTCTTGAGAGGGGCTTCTGAG 529
Db 481 TCTGGTCAGGCCCAAGTCACTGTCCTCTGTCTTGAGAGGGGCTTCTGAG 528

RESULT 4

US-10-027-632-300106
; Sequence 300106, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 300106
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-300106

Query Match 99.7%; Score 527.6; DB 13; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-119;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGTGATGCTGTAGTCGTGCTGATTTATATGCTGATTTATGGTGATTTGCTTCCTT 61
Db 1 ATGGTGATGCTGTAGTCGTGCTGATTTATATGCTGATTTATGGTGATTTGCTTCCTT 60
QY 62 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATA 121
Db 61 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATA 120
QY 122 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAGTTAGCCCTATC 181
Db 121 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAGTTAGCCCTATC 180
QY 182 GGGTATATTTTGGAAAGTTGTAAAACTACTACGTGTTCTCTTCTTAAGTCCCACTCTCTGTT 241
Db 181 GGGTATATTTTGGAAAGTTGTAAAACTACTACGTGTTCTCTTCTTAAGTCCCACTCTCTGTT 240
QY 242 TTCTTTGAGCGAAGAAAGACGATCACCCCTCTCTTCCATGACAAACCAATGACCG 301
|||||

D _b	241	TTCCTTTGACGAGAAAGGAAGCAGTCACTCTCTTCCATGACAACCATGACC	300
Q _y	302	GCTGCTTGCTGTCTCATCACAGGAAGTGCGACAGAGCCCTTTTCTTCCAGGGTTGTG	361
D _b	301	GCTGCTTGCTGTCTCATCACAGGAAGTGCGACAGAGCCCTTTTCTTCCAGGGTTGTG	360
Q _y	362	GGCACCTTGAACTATCTCTGGCTTGCTAGCTAAGAAACTGCCCAAAGTGACGCACACTCAA	421
D _b	361	GGCACCTTGAACTATCTCTGGCTTGCTAGCTAAGAAACTGCCCAAAGTGACGCACACTCAA	420
Q _y	422	GCTGTTGCGGGACATTCCAGAATAAACCAAAACGAGTGATTAATGCTGCTTTGACATCCTC	481
D _b	421	GCTGTTGCGGGACATTCCAGAATAAACCAAAACGAGTGATTAATGCTGCTTTGACATCCTC	480
Q _y	482	TCTGCTCAGGGCCAAGTCATCTGCTCTCTTGGAGAGGGCTTCTGAG	529
D _b	481	TCTGCTCAGGGCCAAGTCATCTGCTCTTGGAGAGGGCTTCTGAG	528

RESULT 5

```

US-10-027-632-54345
; Sequence 54345, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and
; TITLE OF INVENTION: Polymorphisms in
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,000
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,670
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,480
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,211
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,360
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,350
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,000
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4
; SEQ ID NO 54345
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54345

```

Query Match	99.7%	Score	527.6	DB 14	Length	535	
Best Local Similarity	100.0%	Prod. No.	2.6e-119				
Matches	528	Conservative	0	Mismatches	0	Gaps	0
QY	2	ATGGTGATGCTGTAGTCTGGCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTT	61				
Db	1	ATGGTGATGCTGTAGTCTGGCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTT	60				
QY	62	CTTTATACCTTTTATTTATTCCTAAATTTTTCTTAAGCAAAATATTTCTTTTGTCTAATCAATA	121				
Db	61	CTTTATACCTTTTATTTATTCCTAAATTTTTCTTAAGCAAAATATTTCTTTTGTCTAATCAATA	120				
QY	122	AATTATCAAAGGAAAAAACTCGAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATC	181				
Db	121	AATTATCAAAGGAAAAAACTCGAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATC	180				
QY	182	GGGTATATTTTGGAGTTGTAATAATCTAATGTCTCTTCTAAGTCCCACTCTCTGTT	241				
Db	181	GGGTATATTTTGGAGTTGTAATAATCTAATGTCTCTTCTAAGTCCCACTCTCTGTT	240				
QY	242	TTCTTTTGAGCAGGAAGAGAAAGCAGTCACCCCTCTCTTCCATGACAAACCCATGACCG	301				

[illegible]

RESULT. T 6

```

US-10-027-632-300106
;
; SEQUENCE INFORMATION:
;
; APPLICANT: Wang, David G.
;
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
US-10-027-632-300106
; POLYMORPHISMS IN THE HUMAN GENOME
;
; FILE REFERENCE: 108827.129
;
; CURRENT APPLICATION NUMBER: US/10/027,632
;
; CURRENT FILING DATE: 2002-04-30
;
; PRIOR APPLICATION NUMBER: US 60/218,006
;
; PRIOR FILING DATE: 2000-07-12
;
; PRIOR APPLICATION NUMBER: US 60/198,676
;
; PRIOR FILING DATE: 2000-04-20
;
; PRIOR APPLICATION NUMBER: US 60/193,483
;
; PRIOR FILING DATE: 2000-03-29
;
; PRIOR APPLICATION NUMBER: US 60/185,218
;
; PRIOR FILING DATE: 2000-02-24
;
; PRIOR APPLICATION NUMBER: US 60/167,363
;
; PRIOR FILING DATE: 1999-11-23
;
; PRIOR APPLICATION NUMBER: US 60/156,358
;
; PRIOR FILING DATE: 1999-09-28
;
; PRIOR APPLICATION NUMBER: US 60/146,002
;
; PRIOR FILING DATE: 1999-08-09
;
; NUMBER OF SEQ ID NOS: 325720
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 300106
;
; LENGTH: 535
;
; TYPE: DNA
;
; ORGANISM: Human
;
US-10-027-632-300106

```

Query Match	99.7%;	Score 527.6;	DB 14;	Length 535;
Best Local Similarity	100.0%;	Prod. No. 2.6e-119;		
Matches 528; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	2	ATGGTGATGCTGTAGTCGTCGGCTGATTTATATATGCTGATTTATGCTGATTTATGGTGATTTTGGCTTCCTT	61	
Db	1	ATGGTGATGCTGTAGTCGTCGGCTGATTTATATGCTGATTTATGCTGATTTATGGTGATTTTGGCTTCCTT	60	
Qy	62	CTTTATATCTTTTATTTATATCCCAATTTTTCTTTAAGCAAAATATTTCTTTTGTCTAATCAATA	121	
Db	61	CTTTATATCTTTTATTTATTTCCCAATTTTTCTTTAAGCAAAATATTTCTTTTGTCTAATCAATA	120	
Qy	122	AATTTATCAAAAGAAAAAAAACCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGGCCCTATC	181	
Db	121	AATTTATCAAAAGAAAAAAAACCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGGCCCTATC	180	
Qy	182	GGGTATATTTTGGAAAGTTGTAAAAATACTAATGTGTTCTCTCTAAGTCCCACTCCTCTGTGT	241	
Db	181	GGGTATATTTTGGAAAGTTGTAAAAATACTAATGTGTTCTCTCTAAGTCCCACTCCTCTGTGT	240	
Qy	242	TTCTTTTGAGCAGAAAGAGAAAGCAGTCACCCCTCTCTTTCCATGACAAACAACCCCATGACCG	301	

Db 241 TTCTTTGAGCAGGAAGAGACGAGTCACCTCTCTCTTCATGACAAACCCATGACCG 300
QY 302 GCTGCTTGCTGCTCCATCACAGGAAGTGACAGAGCCCTTTTCTTCCAGGGTGTG 361
Db 301 GCTGCTTGCTGCTCCATCACAGGAAGTGACAGAGCCCTTTTCTTCCAGGGTGTG 360
QY 362 GGCACCTTGAACACTCTCGGCTTGCTAGCTAAAGAACTGCCCAAGTGACGACACTCAA 421
Db 361 GGCACCTTGAACACTCTCGGCTTGCTAGCTAAAGAACTGCCCAAGTGACGACACTCAA 420
QY 422 GCTGTTCCGGGACTTCCAGAAATAACCAACAGCTGTATTAAGTCTTTGACATCTC 481
Db 421 GCTGTTCCGGGACTTCCAGAAATAACCAACAGCTGTATTAAGTCTTTGACATCTC 480
QY 482 TCTGCTCAGGCGCAAGTCACTGCTCCTGCTTGGAGAGGGCTTCTGAG 529
Db 481 TCTGCTCAGGCGCAAGTCACTGCTCCTGCTTGGAGAGGGCTTCTGAG 528

RESULT 7

US-09-814-353-17876
; Sequence 17876, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17876
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match 9.7%; Score 51.4; DB 13; Length 321;
Best Local Similarity 59.4%; Pred. No. 0.011;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTTATGGTGATTTGCTTCTTCTTATCTTTATTTATTTCCCAA 86
Db 106 TTTTCTTTAACTTCTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 165
QY 87 TTTTCTTTAAAGCAATATTTCTTGTCTAATCAATAAATATCAAGAAAAAACTGA 146
Db 166 TTTTCTTTTAAATTTTCTTAAATAAAAAAAATTTAAATAAAAAAAATTTT 225
QY 147 AAGCAACGCTTGAAGGAAA 169
Db 226 ANAAAAAAGAAAAAAGAAAAA 248

RESULT 8

US-09-814-353-17553
; Sequence 17553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17553
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-814-353-17553
Query Match 9.6%; Score 51; DB 13; Length 284;
Best Local Similarity 61.8%; Pred. No. 0.013;
Matches 81; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTTATGGTGATTTGCTTCTTCTTATCTTTATTTATTTCCCAA 86
Db 95 TTTTCTTTTAACTTCTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 154
QY 87 TTTTCTTTAAAGCAATATTTCTTGTCTAATCAATAAATATCAAGAAAAAACTGA 146
Db 155 TTTTCTTTTAAATAAAAAATTTTCTTAAATAAAAAAAATTTTAAATAAAAAAA 214
QY 147 AAGCAACGCTT 157
Db 215 AAGAAAGAAAT 225

RESULT 9

US-09-960-352-11218/c
; Sequence 11218, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11218
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 48-LIB3058-026-Q1-K1-D12

US-09-960-352-11218

```
Query Match          9.4%; Score 49.6; DB 10; Length 424;
Best Local Similarity 59.0%; Pred. No. 0.034;
Matches 85; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 26 ATTATATGCTGATTTATGCGTGATTTGCTTCCTCTTTATACCTTTTATTTATCCCAA 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 AATTTAAATTTATTTATTTTAAATTTTATTTATTTTATTTTATTTTATTTTATTTA 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 86 ATTTTCTTAAAGCAATATCTTCTCTATCAATAATATCAAAAGAAAAAAAGCTG 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132 TTTTTTTTTTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTA 73
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 146 AAAGCAACGCTTCAAAAAAGGAA 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 AAAAAAAACCAAAAAA 49
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10

```
US-09-814-353-16036/c
; Sequence 16036, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16036
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-16036
```

```
Query Match          9.1%; Score 48.2; DB 13; Length 316;
Best Local Similarity 63.2%; Pred. No. 0.065;
Matches 74; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 53 TCGTTCCTCTTATCTTTATTTATTTATTTCCCAATTTTCTTAAGCAATATTTCTTGC 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 TGAGTCCCTTTTTCCTTGTTTTCCTCTATTTAGTTTATTTAAATTTTATTTTAA 164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 113 TAATCAATAAATATCAAAAGAAAAAAAGCTGAAAGCAACGCTTGAAAAAGGAAA 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 163 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 107
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 11

```
US-10-311-455-2148/c
; Sequence 2148, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
```

```
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2148
; LENGTH: 113515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2148
```

```
Query Match          9.0%; Score 47.4; DB 13; Length 113515;
Best Local Similarity 54.4%; Pred. No. 2;
Matches 93; Conservative 1; Mismatches 77; Indels 0; Gaps 0;

QY 55 CTTCCTCTCTTATCTTTTATTTATTTCCAAATTTTCTTAAGCAATATTTCTTTGCTA 114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41662 CTTCCTCTCTTATCAATTAATTTACTTTCTATTTTAAACAAACTTTATCTCTT 41603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 115 ATCAATAAATTATCAAAAGAAAAAAAGCTGAAAGCAACGCTTGAAAAAGGAAAGTAG 174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41602 ATTTATATATTATCACTATAAAATCTATAAATTTCTAAATATTTAAATATAATCA 41543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 175 CCCTATCGGTATATTTTGGAGTTGTAATACTACTGTCTCTCTAA 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41542 TTCCCAACATTATTTATTTTAAATATCAATTTATCAAAATTTTAACTTAA 41492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 12

```
US-09-814-353-4872
; Sequence 4872, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4872
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 98, 119, 145, 154, 170, 172, 193, 207, 212, 213, 214,
```


; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18006
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-18006

Query Match 8.8%; Score 46.6; DB 13; Length 383;
Best Local Similarity 56.1%; Pred. No. 0.18; 69; Indels 0; Gaps 0;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 13 GTAGTCGTCGCTGATTATATGCTGATTTATGGGTGATTTGCTTCCTTTTACTTT 72
DB 61 GTGTCGCGCGCGAGTACTTTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTT 120
QY 73 TATTTATCCCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATATCAAAA 132
DB 121 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAATAAAAAA 180
QY 133 GAAAAAACAACGACGACGCTTGAAAAAGGAAA 169
DB 181 AAAAAAATAAAAAAATAAAAAAACAACAAA 217

RESULT 16
US-10-131-827-8192/c
; Sequence 8192, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8192
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-8192

Query Match 8.8%; Score 46.4; DB 12; Length 480;
Best Local Similarity 61.7%; Pred. No. 0.22; 46; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 50 TTTTGCTTCCTCTTTATCTTTTATTTCCCAATTTTCTTAAGCAATATTTCTT 109
DB 142 TTTTACTTTTTTTTCATATTTTAAATATACCAAAATAAATAAATAAATAATTTTA 83
QY 110 TCGTAATCAATAATATCAAGAAAAAATCTGAAAGCAACGCTTGAAAAAGGAAA 169

DB 82 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCTTGTAAAAA 23
RESULT 17
US-09-960-352-8414/c
; Sequence 8414, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8414
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 36-LIB3058-048-Q1-K1-A8
US-09-960-352-8414

Query Match 8.7%; Score 46.2; DB 10; Length 312;
Best Local Similarity 57.1%; Pred. No. 0.2; 63; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 23 CTGATTTATGCTGATTTATGGGTGATTTGCTTCCTTTTATCTTTTATTTATTC 82
DB 192 CTTTTTTTTTTGTT 133
QY 83 CAAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATATCAAGAAAAA 142
DB 132 TTTTTTTTTTTTTTTTTTTTATATATAAAAAAATAAAAAAATAAAAAA 73
QY 143 CTGAAAGCAACGCTTGAAAAAAGGAAA 169
DB 72 AAAAAAAAAAACAACAAAAA 46

RESULT 18
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 8.7%; Score 46; DB 13; Length 3673778;
Best Local Similarity 57.7%; Pred. No. 25; 60; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 23 CTGATTTATGCTGATTTATGGGTGATTTGCTTCCTTTTATCTTTTATTTATTC 82
DB 88260 CCGACTCTTACTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATATTTATCATCT 88201
QY 83 CAAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATATCAAGAAAAA 142

Db 88200 TAAATATTCTCAAAAAAATAATACAAATCATATAATAATAATAAAAAAATC 88141
Qy 143 CTGAAGCAACGCTTGAAAAA 164
Db 88140 AACAAATAACACGTAAACAA 88119

RESULT 19
US-09-814-353-4849
; Sequence 4849, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4849
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 57, 58, 59, 60, 100, 102, 128, 129, 130, 131, 132, 133, 137,
; LOCATION: 138, 140, 141, 142, 143, 150, 174, 176, 177, 184, 185, 190,
; LOCATION: 198, 199, 208, 209, 210, 211, 212, 213, 219, 222, 232, 234,
; LOCATION: 249, 252, 253, 254, 263, 264, 265, 268, 269, 276
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 284, 286, 294, 295, 296, 297, 298, 299, 300, 301, 302, 311,
; LOCATION: 313, 314, 325, 326, 335, 340, 344, 345, 346, 349, 350, 351,
; LOCATION: 359, 367, 379, 382, 386, 387, 389, 394, 395, 403, 406, 408,
; LOCATION: 409, 413, 414, 424, 426, 428, 429, 432, 433, 437, 438
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 439, 440, 441, 442, 444, 447, 451
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4849

Query Match 8.7%; Score 45.8; DB 13; Length 463;
Best Local Similarity 44.9%; Pred. No. 0.3;
Matches 89; Conservative 1; Mismatches 108; Indels 0; Gaps 0;
Qy 67 TACTTTTATTTTCCCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATTA 126
Db 26 TACTTT 85
Qy 127 TCAAAAGAAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTAGCCCTATCGGTA 186
Db 86 TAAAAAATAAAAAAANCNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 145
Qy 187 TATTTTGGAGTTGTAAATACTACTAGTGTCTCTTCTTAAGTCCCACTCCTCTGTTCTT 246

Db 146 TTTTGGGGAATAAAAAAANNNNTTTTTTNNCCCNCCCNATTTTT 205
Qy 247 TGACGAGGAGAGAGAG 264
Db 206 TTNNNNNAAAAAANAANG 223
RESULT 20
US-09-814-353-11146
; Sequence 11146, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11146
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 57, 58, 59, 60, 100, 102, 128, 129, 130, 131, 132, 133, 137,
; LOCATION: 138, 140, 141, 142, 143, 150, 174, 176, 177, 184, 185, 190,
; LOCATION: 198, 199, 208, 209, 210, 211, 212, 213, 219, 222, 232, 234,
; LOCATION: 249, 252, 253, 254, 263, 264, 265, 268, 269, 276
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 284, 286, 294, 295, 296, 297, 298, 299, 300, 301, 302, 311,
; LOCATION: 313, 314, 325, 326, 335, 340, 344, 345, 346, 349, 350, 351,
; LOCATION: 359, 367, 379, 382, 386, 387, 389, 394, 395, 403, 406, 408,
; LOCATION: 409, 413, 414, 424, 426, 428, 429, 432, 433, 437, 438
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 439, 440, 441, 442, 444, 447, 451
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-11146

Query Match 8.7%; Score 45.8; DB 13; Length 463;
Best Local Similarity 44.9%; Pred. No. 0.3;
Matches 89; Conservative 1; Mismatches 108; Indels 0; Gaps 0;
Qy 67 TACTTTTATTTTCCCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATTA 126
Db 26 TACTTT 85
Qy 127 TCAAAAGAAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTAGCCCTATCGGTA 186
Db 86 TAAAAAATAAAAAAANCNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 145
Qy 187 TATTTTGGAGTTGTAAATACTACTAGTGTCTCTTCTTAAGTCCCACTCCTCTGTTCTT 246

[illegible]

```

RESULT 24
US-09-814-353-17876/c
; Sequence 17876, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17876
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: 206..227
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-17876

```

RESULT 25
US-09-814-353-4726/c
; Sequence 4726, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela

/	APPLICANT:	Lillie, James
/	TITLE OF INVENTION:	NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
/	TITLE OF INVENTION:	IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
/	TITLE OF INVENTION:	IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
/	FILE REFERENCE:	MRI-006B
/	CURRENT APPLICATION NUMBER:	US/09/814,353
/	CURRENT FILING DATE:	2001-03-21
/	PRIOR APPLICATION NUMBER:	US 60/191,031
/	PRIOR FILING DATE:	2000-03-21
/	PRIOR APPLICATION NUMBER:	US 60/207,124
/	PRIOR FILING DATE:	2000-05-25
/	PRIOR APPLICATION NUMBER:	US 60/211,940
/	PRIOR FILING DATE:	2000-06-15
/	PRIOR APPLICATION NUMBER:	US 60/216,820
/	PRIOR FILING DATE:	2000-07-07
/	PRIOR APPLICATION NUMBER:	US 60/220,661
/	PRIOR FILING DATE:	2000-07-25
/	PRIOR APPLICATION NUMBER:	US 60/257,672
/	PRIOR FILING DATE:	2000-12-21
/	NUMBER OF SEQ ID NOS:	22037
/	SOFTWARE:	FastSeq for Windows Version 4.0
/	SEQ ID NO 4726	
/	LENGTH:	723
/	TYPE:	DNA
/	ORGANISM:	Homo sapiens
/	FEATURE:	
/	NAME/KEY:	misc feature
/	LOCATION:	56, 57, 59, 137, 147, 148, 156, 162, 167, 169, 173, 189,
/	LOCATION:	200, 202, 214, 223, 225, 227, 231, 239, 240, 248, 257, 259,
/	LOCATION:	260, 266, 270, 273, 274, 275, 277, 278, 280, 283, 287, 295,
/	LOCATION:	306, 311, 313, 317, 340, 344, 351, 353, 361, 367, 370
/	OTHER INFORMATION:	n = A,T,C or G
/	FEATURE:	
/	NAME/KEY:	misc feature
/	LOCATION:	371, 378, 379, 381, 400, 403, 406, 407, 416, 421, 428, 431,
/	LOCATION:	434, 475, 476, 479, 480, 482, 483, 491, 515, 518, 520, 522,
/	LOCATION:	528, 529, 533, 535, 536, 558, 561, 578, 581, 586, 588, 589,
/	LOCATION:	593, 597, 603, 615, 618, 627, 639, 643, 645, 653, 658
/	OTHER INFORMATION:	n = A,T,C or G
/	FEATURE:	
/	NAME/KEY:	misc feature
/	LOCATION:	686, 688, 689, 690, 692, 694, 695, 696, 697, 698, 699, 700,
/	LOCATION:	701, 703, 704, 706, 711, 713, 714
/	OTHER INFORMATION:	n = A,T,C or G
/	US-09-814-353-4726	
/	Query Match	8.5%; Score 44.8; DB 13; Length 723;
/	Best Local Similarity	57.9%; Prod.No 0.67;
/	Matches	70; Conservative 0; Mismatches 51; Indels 0; Gaps
Qy	49	ATTTCCTCCCTGTTTACTTTATTTATTCCTCAAAATTTTCTTAAGCAATAATTTTC
Db	191	AATTTTTTTTTTTTTTTTCCNCNTTTTAAAANAATTTTNNAAAAAAATTTT
Qy	109	TGTGTAATCAATAAATATTCAAAGAAGAAAAAACCTGAACGACCGCTTGAAAAAGACA
Db	131	TTTTTTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA
Qy	169	A 169
Db	71	A 71
RESULT 26		
US-09-814-353-11024/c		
/	Sequence 11024,	Application US/09814353
/	Publication No.	US20030165831A1
/	GENERAL INFORMATION:	
/	APPLICANT:	Lee, John
/	APPLICANT:	Thompson, Pamela
/	APPLICANT:	Lillie, James
/	TITLE OF INVENTION:	NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER

27 TTTATATGCTGATTATGGGTGATTTTGCTTCCTTCTTTATATACTTTTATTTATTCCTCAA 86

Db	686897	TGGTTTGTGCTGCTTTTGATTATCGTTATTTTTTAAAGTTTTTTTTTAGTAGTTGATA	686899
Qy	77	TATTCCCAAAATTTTCTTAAGCAAAATATCTTTGCTAATCAATAAAATATCAAAAGAAA	136
Db	686898	TTTTTAAATAATTTTATGATTAATAATTATATTTTTTATATTTTGAGATAAAATAAAAAATTA	686957
Qy	137	AAAAAACTGAAGCAACGCCTTGA AAAAAGGAAGGTAGCCCTATCGGGTATATTTTGGAA	196
Db	686958	TATAAGCGGTAAAGATAAAATATGAAAAAATAAAAAATATTTTAAATGTTATTAGGAGGATA	687017

Query Match 8.3%; Score 43.8; DB 10; Length 236;

	Query Match	8.2%	Score 43.2	DB 13	Length 397
	Best Local Similarity	54.4%	Pred. No. 1.2		
	Matches 87	Conservative 0	Mismatches 73	Indels 0	Gaps 0
Qy	15	AGTCGGCGTGATTTATATGCTGATTATGGGTGATTTTCGCTTCCTCTCTTTTACTTTT	74		
Db	267	AGTCCTTTGGGATGCTTTTTTAAAGTGTGTGTTTTTCTCTTTATTTTTTTTTTTC	208		
Qy	75	TTTATTTCCCAATTTTTTCTTAAGCAATATTTCTTTGCTAATCAATAAATTATCAAAAGA	134		
Db	207	TTTTTTTTTTTTTAATCCCTTTAAAAAATTTTTTTTTTAAACCTTTAAATAAAAAAGAAAAA	148		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 12:05:58 ; Search time 277 Seconds
(without alignments)

5155.244 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggtgatctgtagtctgt.....tcttgagagggctcttgag 529

Scoring table: IDENTITY.NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : N Geneseq 15Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50.4	9.5	556	23	ABV40063 Human prostate exp
2	50.4	9.5	556	23	ABV40163 Human prostate exp
3	50.4	9.5	556	23	ABV42105 Human prostate exp
4	50.4	9.5	556	23	ABV43601 Human prostate exp
5	49.6	9.4	424	25	ABX46053 Bovine EST associa
6	49.2	9.3	5852	12	AAQ11710 Dictyostelium plas
7	48.6	9.2	464	23	ABV03810 Human prostate exp
8	48.6	9.2	612	22	AAH71471 Human cervical can

9	48	9.1	474	23	ABV56492 Human prostate exp
10	47.8	9.0	300	23	ABV04310 Human prostate exp
11	47.4	9.0	113515	24	ABL34175 Human immune syete
12	47	8.9	309	23	ABV44994 Human prostate exp
13	47	8.9	309	23	ABV44994 Human prostate exp
14	47	8.9	474	23	ABV56492 Human prostate exp
15	47	8.9	612	22	AAH71471 Human cervical can
16	46.6	8.8	312	25	ABX43249 Bovine EST associa
17	46.4	8.8	480	24	ABZ08201 Human leukocyte de
18	46.2	8.7	312	25	ABX43249 Bovine EST associa
19	45.4	8.6	556	23	ABV40063 Human prostate exp
20	45.4	8.6	556	23	ABV40163 Human prostate exp
21	45.4	8.6	556	23	ABV42105 Human prostate exp
22	45.4	8.6	556	23	ABV43601 Human prostate exp
23	44.6	8.4	464	23	ABV03810 Human prostate exp
24	44.6	8.4	10988	24	ABL70200 Chemically treated
25	44.6	8.4	10988	24	AA561158 Human gene regulat
26	44.6	8.4	10988	24	AA561158 Signal transductio
27	44.4	8.4	425	22	AA560450 Human cancer agent
28	44.4	8.4	494	23	ABV10021 Human prostate exp
29	44.4	8.4	664	21	AAA26336 Human secreted pro
30	44.4	8.4	6078	22	AAA26406 Tumour suppressor
31	44.4	8.4	6078	24	ABL33137 Human immune syete
32	43.8	8.3	198	22	AA507727 Cervical cancer pr
33	43.8	8.3	236	25	ABX47018 Bovine EST associa
34	43.8	8.3	323	23	ABV60996 Human prostate exp
35	43.8	8.3	378	22	AAI90863 Human polynucleoti
36	43.8	8.3	411	22	AAI88595 Human polynucleoti
37	43.8	8.3	9884	24	ABK33925 Human DNA for stag
38	43.8	8.3	56153	22	AA546793 Tumour suppressor
39	43.6	8.2	626	23	ABV60941 Human prostate exp
40	43.6	8.2	6590	24	ABL33476 Human immune syete
41	43.4	8.2	300	22	AAH70049 Human cervical can
42	43.4	8.2	424	25	ABV46053 Bovine EST associa
43	43.4	8.2	431	23	ABV04399 Human prostate exp
44	43.2	8.2	380	25	ABX44170 Bovine EST associa
45	43.2	8.2	430	24	ABV83644 Human breast spec
46	43.2	8.2	12426	22	AA546481 Tumour suppressor
47	43	8.1	255	22	AAH82206 Rat differential t
48	43	8.1	180557	24	ABN85750 Human BAC clone Rp
49	43	8.1	368004	24	ABN85750 Human transporter
50	42.8	8.1	404	23	ABV56394 Human prostate exp
51	42.6	8.1	522	22	AA534084 Human cDNA encodin
52	42.6	8.1	621	22	AAH71551 Human cervical can
53	42.6	8.1	768	22	AAAL08093 Human breast cance
54	42.4	8.0	308	25	ABX42505 Bovine EST associa
55	42.4	8.0	445	23	ABV05431 Human prostate exp
56	42.4	8.0	524	25	ABT22885 Breast cancer mark
57	42.4	8.0	2672	19	AAV39862 Mycoplasma hyopneu
58	42.4	8.0	6731	24	ABL33060 Human immune syete
59	42.4	8.0	277	25	ABX47508 Bovine EST associa
60	42.2	8.0	277	25	ABX47508 Bovine EST associa
61	42.2	8.0	294	22	AAH70082 Human cervical can
62	42.2	8.0	380	25	ABX44170 Bovine EST associa
63	42.2	8.0	397	22	AAI82664 Human polynucleoti
64	42.2	8.0	411	22	AAI88595 Human polynucleoti
65	42.2	8.0	418	25	ABX39680 Bovine EST associa
66	42.2	8.0	453	23	ABV09564 Human prostate exp
67	42.2	8.0	493	22	AAAL25179 Human breast cance
68	42.2	8.0	550	23	ABV57072 Human prostate exp
69	42.2	8.0	1066	24	ABN98373 Arabidopsis thalia
70	42.2	8.0	5359	24	ABL33607 Human immune syete
71	42.2	8.0	5359	24	ABK28272 DNA transcrip
72	42.2	8.0	6533	24	ABK28383 DNA transcrip
73	42	7.9	1342	21	AAA26424 Human secreted pro
74	42	7.9	4163	24	ABL34330 Human immune syete
75	42	7.9	13449	22	ABL33384 Human immune syete
76	41.8	7.9	389	22	AAAL20088 Human breast cance
77	41.8	7.9	396	22	AAF94842 Human ovarian can
78	41.8	7.9	396	24	ABT03109 Human ovarian carc
79	41.8	7.9	396	24	ABL48792 Ovarian carcinoma
80	41.8	7.9	612	22	AAH71474 Human cervical can
81	41.8	7.9	2001	18	AAAT59700 PTH-like peptide D

C 82 41.8 7.9 5020 25 ABZ09960
 C 83 41.8 7.9 513445 22 AAI61373
 C 84 41.6 7.9 198 22 AAS07727
 C 85 41.6 7.9 298 25 ABX35839
 C 86 41.6 7.9 545 22 AAH70126
 C 87 41.6 7.9 566 20 AAY86948
 C 88 41.6 7.9 693 6 AAN50336
 C 89 41.6 7.9 693 11 AAQ06408
 C 90 41.6 7.9 1180 24 AAD33724
 C 91 41.6 7.9 6794 24 ABL70220
 C 92 41.6 7.9 6794 24 AAS61175
 C 93 41.6 7.9 6794 24 ABK31265
 C 94 41.6 7.9 22073 22 AAL04663
 C 95 41.6 7.9 22073 23 ABL97570
 C 96 41.6 7.9 22073 25 ABZ74560
 C 97 41.6 7.9 22073 25 ABZ68092
 C 98 41.4 7.8 418 25 ABX39680
 C 99 41.4 7.8 643 22 AAL16045
 C 100 41.4 7.8 1617 24 ABS59408

ALIGNMENTS

RESULT 1
 ABV40063
 ID ABV40063 standard; cDNA; 556 BP.
 AC ABV40063;
 DT 16-SEP-2002 (first entry)
 XX
 XX
 XX
 DE Human prostate expression marker cDNA 40054.
 DE
 DE
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 XX
 PN WO200160860-A2.
 XX
 XX
 PD 23-AUG-2001.
 XX
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 XX
 DR WPI; 2001-662795/76.
 XX
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 XX
 PS Claim 1; Page 8100; 11750pp; English.
 XX
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynanamic or pharmacogenomic marker.
 XX
 SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;
 Query Match 9.5%; Score 50.4; DB 23; Length 556;
 Best Local Similarity 58.8%; Pred. No. 0.02;
 Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 22 GCTGATTTATATGCTGATTATGCGTGGATTTCCTTCTTATCTTTATTTATTC 81
 Db 1 GCTTTTATTTTTCCTTCTTATCTTTATTTTTCCTTCTTATCTTTATTTATTC 60
 QY 82 CCAATTTTCTTAAGCAATAATTTCTTGTGTAATCAATAAATATCAAGAAAAAAA 141
 Db 61 TTTTTCCTTCTTATCTTTATCTTTATTTTTCCTTCTTATCTTTATTTATTC 120
 QY 142 ACTGAAGCAACGCTTGAAGAAAGAAA 169
 Db 121 AAAAAAAGAAAAAAGAAAAAAGAAAAA 148

RESULT 2
 ABV40163
 ID ABV40163 standard; cDNA; 556 BP.
 AC ABV40163;
 DT 16-SEP-2002 (first entry)
 XX
 XX
 DE Human prostate expression marker cDNA 40154.
 DE
 DE
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 XX
 PN WO200160860-A2.
 XX
 XX
 PD 23-AUG-2001.
 XX
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 XX
 DR WPI; 2001-662795/76.
 XX
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 XX
 PS Claim 1; Page 8115; 11750pp; English.
 XX
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;

Query Match 9.5%; Score 50.4; DB 23; Length 556;
Best Local Similarity 58.8%; Pred. No. 0.02;
Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 GCTGATTTATGCTGATTTATGCTGATTTGCTTCTTATATCTTTATTTATTTTC 81
Db 1 GCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60

Qy 82 CCAATTTTCTTAAGCAAAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAA 141
Db 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120

Qy 142 ACTGAAGCAACGCTTGAAAAAGAAA 169
Db 121 AAAAAAAGAAAAAAGAAAAAAGAAAAA 148

RESULT 3
ABV42105
ID ABV42105 standard; cDNA; 556 BP.
XX
AC ABV42105;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 42096.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 8441; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;

Query Match 9.5%; Score 50.4; DB 23; Length 556;
Best Local Similarity 58.8%; Pred. No. 0.02;
Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 GCTGATTTATGCTGATTTATGCTGATTTGCTTCTTATATCTTTATTTATTTTC 81
Db 1 GCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60

Qy 82 CCAATTTTCTTAAGCAAAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAA 141
Db 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120

Qy 142 ACTGAAGCAACGCTTGAAAAAGAAA 169
Db 121 AAAAAAAGAAAAAAGAAAAAAGAAAAA 148

RESULT 4
ABV43601
ID ABV43601 standard; cDNA; 556 BP.
XX
AC ABV43601;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 43592.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 8681-8682; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;
 Query Match 9.5%; Score 50.4; DB 23; Length 556;
 Best Local Similarity 58.8%; Pred. No. 0.02; Indels 0; Gaps 0;
 Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 22 GCGATTATATGCTGATTATGCGTATGCTTCCTCTTATATCTTTATTTATTC 81
 DB 1 GCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60
 QY 82 CCAATTTCTTAAGCAATATCTTCTGCTAATCAATTAATCAAAAGAAAAA 141
 DB 61 TTTTCTTTTCTTAAATTTTCTTTTCTTAAATTAATAAAAAA 120
 QY 142 ACTGAAGCAAGCTTGAAAAAGGAAA 169
 DB 121 AAAAAA 148
 RESULT 5
 ABX46053/c
 ID ABX46053 standard; cDNA; 424 BP.
 XX
 AC ABX46053;
 DT 21-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #11218.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-0960352.
 XX
 PR 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX
 DR New nucleic acid associated with lactation, and muscle and fat
 XX deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 PT
 PS Claim 2; SEQ ID No 11218; 245pp; English.
 XX
 XX The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMFD), derived
 CC from cattle, and the LMFD nucleic acid can specifically hybridize to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are; (1) a transformed cell having a nucleic acid
 CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMFD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMFD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139.
 XX
 SQ Sequence 424 BP; 228 A; 7 C; 32 G; 157 T; 0 other;
 Query Match 9.4%; Score 49.6; DB 25; Length 424;
 Best Local Similarity 59.0%; Pred. No. 0.028;
 Matches 85; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
 QY 26 ATTTATATGCTGATTATGCGTATGCTTCCTCTTATCTTTATTTATTTCCCA 85
 DB 192 AATTTTAAATTTATTTATTTTAAATTTTATTTATTTATTTATTTATTTT 133
 QY 86 ATTTTCTTAAAGCAATATTTCTTTGCTAATCAATAAATATCAAAAGAAAAA 145
 DB 132 TTTTCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 133
 QY 146 AAAGCAAGCTTGAAAAAGGAAA 169
 DB 72 AAAAAA 49
 RESULT 6
 AAQ11710/c
 ID AAQ11710 standard; DNA; 5852 BP.
 XX
 AC AAQ11710;
 XX
 DT 25-MAR-2003 (updated)
 DT 30-JUL-1991 (first entry)
 XX
 DE Dictyostelium plasmid Ddp2 containing Rep gene.
 XX
 KW slime mould; replication; Rep gene; ss.
 XX
 OS Dictyostelium discoideum.
 XX
 FH Key Location/Qualifiers
 FT CDS 2378..5041
 FT /*tag= a
 FT /product= involved in extrachromosomal replication
 XX
 XX WO9106644-A.
 XX
 XX 16-MAY-1991.
 XX
 XX 02-NOV-1990; 90WO-AU00530.
 XX
 XX 02-NOV-1989; 89AU-0007187.

XX PA (UTMA-) UNIV MACQUARIE.
 XX PI Slade MB, Chang ACM, Williams KL;
 XX DR WPI; 1991-164194/22.
 XX DR P-PSDB; AAR11988.
 XX PT Polypeptide facilitating extra-chromosomal replication - of
 XX PT recombinant plasmid in Dictyostelium species
 XX PS Claim 15; Fig 1; 90pp; English.
 XX CC The sequence of Ddp2 has been found to contain the putative open
 CC reading frame indicated in the Features Table. The possible ORF is
 CC flanked by regions with similarity to promoter and poly adenylation
 CC signals of known Dictyostelium genes. The RNA and polypeptide
 CC product of the Rep gene have not, however, been detected. It is
 CC believed that the product is produced in low amounts to positively
 CC regulate initiation of plasmid replication. The polypeptide may also
 CC contain regions that act as negative regulators of plasmid copy
 CC number. See also AAQ11711 and AAQ11712.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 5852 BP; 2298 A; 651 C; 708 G; 2195 T; 0 other;
 Query Match 9.3%; Score 49.2; DB 12; Length 5852;
 Best Local Similarity 57.0%; Pred. No. 0.076; 68; Indels 0; Gaps 0;
 Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 12 TGTAGTCGTGCTGATTATATGCTGATTATGCGTGAATTTGCTTCTCTTATACCTT 71
 DB 5767 TTTAATTATTTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTCTTACTT 5708
 QY 72 TTAITATTCCTCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATATCAAA 131
 DB 5707 ATATTTCTATTTTATTTATTAATAATTAATTAATTAATTAATAATAAAAAA 5648
 QY 132 AGAAAAAACTGAAGCAAGCTTGAAAAAGGAAA 169
 DB 5647 AAAAAAAATTTAATTAATAATAATAATAATAATAAAAAA 5610
 RESULT 7
 ABV03810
 ID ABV03810 standard; cDNA; 464 BP.
 XX AC ABV03810;
 XX DT 13-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 3801.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US05171.
 XX PR 17-FEB-2000; 2000US-183319P.
 XX PR 16-MAR-2000; 2000US-189862P.
 XX PR 25-MAY-2000; 2000US-207454P.
 XX PR 09-JUN-2000; 2000US-211314P.
 XX PR 18-JUL-2000; 2000US-219007P.
 XX PR 13-DEC-2000; 2000US-255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX

PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.
 XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 XX PT prostate cells and correlating with presence of prostate cancer, useful
 XX PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX PS Claim 1; Page 677; 11750pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX SQ Sequence 464 BP; 199 A; 74 C; 16 G; 116 T; 59 other;
 Query Match 9.2%; Score 48.6; DB 23; Length 464;
 Best Local Similarity 43.3%; Pred. No. 0.05; 172; Indels 0; Gaps 0;
 Matches 132; Conservative 1; Mismatches 172; Indels 0; Gaps 0;
 QY 27 TTTATATGCTGATTATGCGGATTTTGCTTCTCTTATCTTTTATTTATTTCCCAA 86
 DB 8 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 67
 QY 87 TTTTCTTAAGCAATATTTCTTTGCTAATCAATAATTAATCAAAAGAAAAAACTGA 146
 DB 68 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
 QY 147 AAGCAAGCTTGAAAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAT 206
 DB 128 AAAAAAAATTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAA 187
 QY 207 ACTACRTGTTCTCTTAAGTCCCACTCTCTGTTTCTTTGAGCAGAGAGAGAGCA 266
 DB 188 AAAAAAAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAA 247
 QY 267 GTCACCTCTCTCTCCATGACAAACCCATGACCGGCTGCTGCTGCTCATCAGGA 326
 DB 248 AACCCCNCCANNNNNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 307
 QY 327 AGGTG 331
 DB 308 AGGGG 312
 RESULT 8
 AAH71471
 ID AAH71471 standard; cDNA; 612 BP.
 XX AC AAH71471;
 XX DT 19-SEP-2001 (first entry)
 XX DE Human cervical cancer marker nucleic acid 2745.
 XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN WO200142467-A2.
 XX

```

PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
XX 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 564; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
XX Sequence 612 BP; 380 A; 50 C; 2 G; 178 T; 2 other;
SQ
Query Match 9.2%; Score 48.6; DB 22; Length 612;
Best Local Similarity 58.7%; Pred. No. 0.054;
Matches 84; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 27 TTATATGCTGATTTATGGTGATTTGCTTCCTCTTATCTTTATTTATTTATTTCCCAA 86
Db 17 TTTTCTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 76
QY 87 TTTTCTTAAAGCAATATTTCTTTGCTAATCAATAATATCAAAAGAAAAAACTGCA 146
Db 77 TTTTCTTTTAAATTTTCTTTTAAATTAATAAAAAAAATTAATAAAAAAAATTAATA 136
QY 147 AAGCAAGCTTGAAGAAAGAAA 169
Db 137 AAAAAAAATAAAAAAAATTAATA 159
RESULT 9
ABV56492
ID ABV56492 standard; cDNA; 474 BP.
XX
XX ABV56492;
XX
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 56483.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR
25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JB;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 10898; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynanamic or pharmacogenomic marker.
XX
XX Sequence 474 BP; 234 A; 34 C; 107 G; 98 T; 1 other;
SQ
Query Match 9.1%; Score 48; DB 23; Length 474;
Best Local Similarity 59.6%; Pred. No. 0.07;
Matches 81; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 34 GCTGATTTATGGTGATTTGCTTCCTCTTATCTTTATCTTTATTTATTTCCCAATTTTCT 93
Db 9 GCGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 68
QY 94 TAAGCAAAATATTTCTTCTCAATCAATAAATATCAAAAGAAAAAACTCAAGCAAC 153
Db 69 TTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTAATAAAAAAAATTAATAAAAAAA 128
QY 154 GCTTCAAAAAAGGAAA 169
Db 129 AAAAAAAATAAAAAAATTAATA 144
RESULT 10
ABV04310/c
ID ABV04310 standard; cDNA; 300 BP.
XX
XX ABV04310;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 4301.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX

```


PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JB;
 XX
 XX WPI; 2001-662795/76.
 DR
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 750; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 XX Sequence 300 BP; 125 A; 22 C; 19 G; 101 T; 33 other;
 SQ
 Query Match 9.0%; Score 47.8; DB 23; Length 300;
 Best Local Similarity 56.2%; Pred. No. 0.068;
 Matches 82; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 24 TGATTATATGCTGATTTATGGGTGATTTGCTCTTTATATCTTTATTTATTTCC 83
 Db 168 TGAATAAATTTNNTTTTTTTGGAAATTTNNTTTTTTTTCCTTTTAAATTTTTT 109
 QY 84 AAATTTTCTTAAGCAATATTTCTTGTCTATCAATCAATATCAAGRAAARAAAC 143
 Db 108 TNNTTTTTAAATTAATTTTTTTTTTTTTCNCCAAAAAATAAAAAAAAAAAAAA 49
 QY 144 TGAAGCAACGCTTGAAAAAGGAAA 169
 Db 48 AAAAAAGAAAAAATAAAAAAAAAA 23
 RESULT 11
 ABL34175/c
 ID ABL34175 standard; DNA; 113515 BP.
 AC ABL34175;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 DE Human immune system associated gene SEQ ID NO: 2148.
 DE
 XX Human; immune system disease; cytosine methylation; antiaesthatic;
 KW antiaeriosclerotic; antianaemic; cytosatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritis; antidiabetic; antipsoriasis;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 XX gene; ds.
 XX

OS Homo sapiens.
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 2148; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 113515 BP; 35172 A; 1174 C; 25250 G; 54649 T; 0 other;
 Query Match 9.0%; Score 47.4; DB 24; Length 113515;
 Best Local Similarity 54.4%; Pred. No. 0.49;
 Matches 93; Conservative 1; Mismatches 77; Indels 0; Gaps 0;
 QY 55 CTTCCTCTTTATCTTTTATTTATTTCCCAATTTTCTTAAGCAATATTTCTTTGCTA 114
 Db 41662 CTTCCTCTTTATCAATTAATAATTTACTTTCTTTTAAACAACTTTATCTCTT 41603
 QY 115 ATCAATAAATTTATCAAGAAAAAATACTGAAAGCAACGCTTGAAAAAGGAAAGTTAG 174
 Db 41602 ATTATATATTTACTATAAATACTATAAATCTATAAATAATAAATAAATAAATCA 41543
 QY 175 CCTATCGGGTATATTTTGGAAAGTTGTAAATACTACTGTCTCTCTAA 225
 Db 41542 TTCCACCATTATTTATTTTAAATATCAATTTATCAAAATTTAACTTATA 41492
 RESULT 12
 ABV44994
 ID ABV44994 standard; cDNA; 309 BP.
 XX
 AC ABV44994;
 XX
 XX 16-SEP-2002 (first entry)
 DT
 XX
 DE Human prostate expression marker cDNA 44985.
 DE
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 PD
 XX 20-FEB-2001; 2001WO-US05171.
 PF
 XX 17-FEB-2000; 2000US-183319P.
 PR

PR 11-JAN-2000; 2000US-0480902.
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2; SEQ ID NO 8414; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34835-ABX49947, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=200201371139.
XX
XX Sequence 312 BP; 167 A; 27 C; 28 G; 90 T; 0 other;
XX
XX Query Match 8.8%; Score 46.6; DB 25; Length 312;
Best Local Similarity 57.0%; Pred. No. 0.13;
Matches 85; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
XX
QY 21 GGCTGATTTATGCTGATTGTTGGGTGATTTGGCTTCCTTTATATCTTTATTTATT 80
Db 25 GCTCAATTT 84
XX
QY 81 CCCAAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAA 140
Db 85 TTAAATTTTTTTTTTTTTTTTATTAATAATAAAAAAATAAAAAAATAAAAAA 144
XX
QY 141 AACTGAAGCAAGCTTTGAAAAAGGAAA 169
Db 145 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 173
XX
RESULT 17
ABZ08201/c
ID ABZ08201 standard; cDNA; 480 BP.
XX
XX ABZ08201;
XX
XX 09-JAN-2003 (first entry)
DT
DE Human leukocyte derived cDNA SEQ ID NO 8192.

XX Human; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; gene;
KW ss.
XX
XX Homo sapiens.
XX
XX WO200257414-A2.
XX
XX 25-JUL-2002.
XX
XX 22-OCT-2001; 2001WO-US47856.
XX
XX 20-OCT-2000; 2000US-241994P.
XX
XX 08-JUN-2001; 2001US-296764P.
XX
XX (BIOC-) BIOCARDIA INC.
XX
XX Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quettermous T, Johnson F;
XX WPI; 2002-636525/68.
XX
XX New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides
XX
XX Claim 26; Page 1829-1830; 2038pp; English.
XX
XX The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08132) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC predicting disease complications in an individual or monitoring response
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection. The
CC present sequence is that of a human leukocyte expressed cDNA of the
CC invention.
XX
XX Sequence 480 BP; 124 A; 34 C; 159 G; 163 T; 0 other;
XX
XX Query Match 8.8%; Score 46.4; DB 24; Length 480;
Best Local Similarity 61.7%; Pred. No. 0.17;
Matches 74; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
XX
QY 50 TTTTGCTTCCTTCTTTATCTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 109
Db 142 TTTTACTTTTTTTTTTTCATATTTTTTAAATACCAAAATAATAAAATAAAATATTTTA 83
XX
QY 110 TGCTAATCAATAAATTTATCAAAAGAAAAAATAAAAGCAACGCTTGAAAAAGGAAA 169
Db 82 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 23
XX
RESULT 18
ABX43249/c
ID ABX43249 standard; cDNA; 312 BP.
XX
XX ABX43249;
XX
XX 20-FEB-2003 (first entry)
DT
XX Bovine EST associated with lactation/muscle/fat deposition #8414.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.


```
Db 56 AAAAAAAAAAAAAAAAAAAAAA 34
RESULT 20
ABV40163/c
ID ABV40163 standard; cDNA; 556 BP.
XX AC ABV40163;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 40154.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer -
XX PS Claim 1; Page 8115; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX CC cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX CC in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a
XX CC patient;
XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;
Query Match 8.6%; Score 45.4; DB 23; Length 556;
Best Local Similarity 57.3%; Pred. No. 0.3;
Matches 82; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 27 TTATATGCTGATTTAAGGGTGGTTCCTCTTTATCTTTTATCTTTTATTTATTTATCCCAA 86
Db 176 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 117
QY 87 TTTTCTTTAAGCAATATTTCTTGTCTATCAATTAATATCAAGAAAAAACTGA 146
Db 116 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 57
us-10-027-632-1.rng
QY 147 AAGCAACGCTTCAAAAAAGGAAA 169
Db 56 AAAAAAAAAAAAAAAAAAAAAA 34
RESULT 21
ABV42105/c
ID ABV42105 standard; cDNA; 556 BP.
XX AC ABV42105;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 42096.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer -
XX PS Claim 1; Page 8441; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX CC cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX CC in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a
XX CC patient;
XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;
Query Match 8.6%; Score 45.4; DB 23; Length 556;
Best Local Similarity 57.3%; Pred. No. 0.3;
Matches 82; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 27 TTATATGCTGATTTAAGGGTGGTTCCTCTTTATCTTTTATCTTTTATTTATTTATCCCAA 86
Db 176 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 117
QY 87 TTTTCTTTAAGCAATATTTCTTGTCTATCAATTAATATCAAGAAAAAACTGA 146
```


CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX
SQ Sequence 10988 BP; 3128 A; 225 C; 2649 G; 4986 T; 0 other;

Query Match 8.4%; Score 44.6; DB 24; Length 10988;
Best Local Similarity 60.2%; Pred. No. 1,1;
Matches 74; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 12 TGTAGTCGTGGCTGATTTATATGCTGATTTATGGGTGATTTTGGCTCTTTATACCT 71
10263 TGTAGATGGTGGTGATATATTTTTGGTTTTATATTTTTTTTATAGTATACGTTT 10322

QY 72 TTATTTATCCCAATTTTTCTTAGCAATATTTCTTTGCTATCATATAATATCAAA 131
10323 TTTTATATTTTAAATGGATTTTGAATTTTATTTATAGAAATTTAAAGTATTATCGAT 10382

QY 132 AGA 134
Db 10383 AAA 10385

RESULT 27
AAS60450
ID AAS60450 standard; cDNA; 425 BP.
XX
AC AAS60450;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human cancer agent-sensitive marker #181.
XX
KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmacytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
OS Homo sapiens.
XX
PN WO200179556-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12132.
XX
PR 14-APR-2000; 2000US-197538P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Brown JL, Bolt A, Van Huffel C;
XX
DR WPI; 2001-602933/68.
XX
PT Novel nucleic acid, used as a marker to determine the effectiveness of
XX using TAXOL to treat cancer cell growth in individuals -
XX
XX Claim 1; Page 221; 527pp; English.
XX
CC The invention relates to 1046 novel nucleic acids which are used as
CC markers for determining the sensitivity of a cancer cell to the
CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
CC they are shown to express one of the 242 sensitivity markers or the
CC cells are shown not to express one of the 804 resistance markers.
CC The methods can be used to determine the effectiveness of TAXOL
CC in the treatment of cancer cell growth in an individual. The markers
CC can be used as targets in developing anti-cancer agents such as
CC chemotherapeutic compounds. The markers can also be used as targets in
CC developing treatments for cancer, particularly those cancers which

KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antihaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 1110; 32pp + Sequence Listing; German.

XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 6078 BP; 1615 A; 218 C; 1499 G; 2746 T; 0 other;

Query Match 8.4%; Score 44.4; DB 24; Length 6078;
 Best Local Similarity 54.4%; Pred. No. 1.1;
 Matches 87; Conservative 1; Mismatches 72; Indels 0; Gaps 0;
 QY 65 TATACCTTTTATTTCCCAAAATTTTCTTAAGCAAAATATTTCTTGTAAATCAATAAT 124
 DB 4494 TAAATATATACCAAAATCTTAAATATATATAAAAAAAACCTCACACATAAATAAAA 4435
 QY 125 TATCAAAAGAAAAAAACCTGAAAGCAACGCTGAAAAAGGAAGTTAGCCCTATCGGG 184
 DB 4434 AAAAAAAGAAAAAACAATATTAACCAAACTCTTCAATATTTATCTTAAATACCTTTT 4375
 QY 185 TATATTTTGGAGCTGTAAATACTACRTGTTCTCTCTTA 224
 DB 4374 TATTTTATATTTTATATATTTTCCAAATTTCTACAA 4335

RESULT 32
 ID AAS07727
 XX AAS07727 standard; DNA; 198 BP.
 AC AAS07727;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Cervical cancer pre-malignant condition DNA marker #26.
 XX
 KW Cervical cancer; pre-malignant condition marker; tumorigenesis; CIN; ds;
 KW cervical intraepithelial neoplasia; squamous intraepithelial lesions;
 KW polymerase chain reaction; PCR; enzyme linked immunosorbent assay; ELISA;

KW anticancer therapy; carcinogen; antisense inhibition.
 XX
 OS Homo sapiens.
 XX
 PN WO200142792-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WO-US33311.
 XX
 PR 08-DEC-1999; 99US-0169811.
 PR 21-DEC-1999; 99US-0171330.
 PR 14-MAR-2000; 2000US-0189113.
 PR 31-MAR-2000; 2000US-0193943.
 PR 12-MAY-2000; 2000US-0203772.
 PR 09-JUN-2000; 2000US-0210820.
 PR 21-JUL-2000; 2000US-0220113.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX
 DR WPI; 2001-367889/38.

XX
 CC Cervical cancer protein markers useful for the diagnosis, prevention
 CC and treatment of cervical cancers, especially cervical intraepithelial
 CC neoplasia or squamous intraepithelial lesions -
 XX
 PS 20; Page 407; 436pp; English.
 XX
 CC The sequence represents a cervical cancer pre-malignant condition marker.
 CC These markers encode proteins that are over expressed during
 CC tumorigenesis. The proteins and their corresponding nucleic acid
 CC sequences can therefore be used for the diagnosis, prevention and
 CC treatment of cervical cancers, particularly cervical intraepithelial
 CC neoplasia (CIN) and squamous intraepithelial lesions (SIL). The sequences
 CC may be used as markers in diagnostic assays to detect cancerous
 CC conditions (e.g. by polymerase chain reaction (PCR) or enzyme linked
 CC immunosorbent assay (ELISA)), to monitor the efficacy of anticancer
 CC therapies and to identify anticancer or carcinogenic compounds. The level
 CC of expression of a marker in a patient sample is compared with the normal
 CC level of expression of the marker in a control non-cervical cancer
 CC sample, whereby a significant difference indicates that the patient is
 CC afflicted with cervical cancer or a pre-malignant condition. The
 CC expression of these proteins may be inhibited by antisense inhibition for
 CC the treatment of cancers. They may be used in this way for the treatment
 CC of cervical intraepithelial neoplasia or squamous intraepithelial
 CC lesions.
 XX
 SQ Sequence 198 BP; 67 A; 11 C; 10 G; 102 T; 8 other;

Query Match 8.3%; Score 43.8; DB 22; Length 198;
 Best Local Similarity 54.5%; Pred. No. 0.53;
 Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 QY 27 TTTATATGCTGATTTATGGTGATTTTGTCTCTCTTTTATCTTTTATTTATTTTCCCAA 86
 DB 40 TTTTTCCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 99
 QY 87 TTTTTCCTTAAGCAAAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACTGCA 146
 DB 100 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 159
 QY 147 RAGCAACGCTGAAAAAGGAAA 169
 DB 160 AANANNAAAAAAANAAAAA 182

RESULT 33
 ABX47018
 ID ABX47018 standard; cDNA; 236 BP.
 XX
 AC ABX47018;

[illegible]

DT XX 21-FEB-2003 (first entry)

DE XX Bovine EST associated with lactation/muscle/fat deposition #9335.

KW XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

KW XX muscle deposition; fat deposition; genome mapping; gene identification;

KW XX gene analysis; cattle breeding.

OS XX Bos Taurus.

XX XX

PN XX US2002137139-A1.

XX XX

PD XX 26-SEP-2002.

XX XX

PF XX 24-SEP-2001; 2001US-0960352.

XX XX

PR XX 12-JAN-1999; 99US-115707P.

PR XX 11-JAN-2000; 2000US-0480902.

XX XX

PA XX (BYAT/) BYATT J C.

PA XX (MATH/) MATHIALAGAN N.

PA XX (TAON/) TAO N.

PA XX (WARR/) WARREN W C.

XX XX

PI XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX XX WPI; 2003-110599/10.

DR XX

XX XX

PT XX New nucleic acid associated with lactation, and muscle and fat

PT XX deposition, useful for genome mapping, gene identification and

PT XX analysis, cattle breeding, or for genetically improving cattle

XX XX

PS XX Claim 2; SEQ ID No 9335; 245pp; English.

XX XX

CC XX The invention relates to a purified nucleic acid molecule associated with

CC XX lactation or muscle and fat deposition (designated LMFD), derived

CC XX from cattle, and the LMFD nucleic acid can specifically hybridize to a

CC XX second nucleic acid molecule comprising any of 15112 nucleotide

CC XX sequences, appearing as ABX34836-ABX49947, or complements of them.

CC XX Also included are: (1) a transformed cell having a nucleic acid

CC XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-

CC XX translated sequence that functions in the cell to cause termination of

CC XX transcription and addition of polyadenylated ribonucleotides to a 3' and

CC XX of the mRNA molecule; and (2) determining a level or pattern of a

CC XX molecule in a bovine cell or tissue comprising: (a) incubating a marker

CC XX nucleic acid (comprising any of the 15112 nucleic acid sequences or its

CC XX complement or fragment) with a complementary nucleic acid molecule

CC XX obtained from the bovine cell or tissue, where hybridisation between the

CC XX marker nucleic acid and the complementary nucleic acid permits the

CC XX detection of the molecule; and (b) detecting the level or pattern of the

CC XX complementary nucleic acid, where the detection of the complementary

CC XX nucleic acid is predictive of the level or pattern of the molecule.

CC XX The LMFD nucleic acid is used for determining a level or pattern

CC XX of a molecule in a bovine cell or tissue. It is useful for genome

CC XX mapping, gene identification and analysis, cattle breeding, preparation

CC XX of constructs for use in cattle gene expression, or for genetically

CC XX improving cattle. The present sequence is one of the 15112 bovine

CC XX LMFD EST (expressed sequence tag) nucleic acids

CC XX

CC XX Note: The present sequence was not shown in the specification but

CC XX was obtained in electronic format from the USPTO web site:

CC XX seqdata.uspto.gov/sequence.html?DocID=20020137139.

XX XX

SQ XX Sequence 380 BP; 154 A; 22 C; 46 G; 158 T; 0 other;

Query Match 8.2%; Score 43.2; DB 25; Length 380;

Best Local Similarity 56.2%; Pred. No. 0.9;

Matches 81; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTATGGTGATTTGCTTCCTCTTATATCTTTATTTATCCCAA 86

DB 49 TTTTGGTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 108

QY 87 TTTTCTTAAGCAAAATATTTCTTCTGTAATCAATAATATCAAAAGAAAAAACTGA 146

DB 109 TTTTCTTATTTATAATAATATTTTCAAAATTTAAAAAATAATAATGTTGTTAAAAAATTTT 168

QY 147 AAGCAACGCTTGAAAAAGGAAG 170

DB 169 AAAAAAATTTAAAAAAGAAAAAG 192

RESULT 45

ABV83644

ID ABV83644 standard; cDNA; 430 BP.

XX XX

AC ABV83644;

XX XX

DT 06-DEC-2002 (first entry)

XX XX

DE Human breast specific gene SEQ ID NO 87.

XX XX

KW Human; breast; cytostatic; cancer; transgenic; gene therapy; vaccine;

KW KW gene; ss.

XX XX

OS Homo sapiens.

XX XX

FN WO200266605-A2.

XX XX

PD 29-AUG-2002.

XX XX

PF 14-FEB-2002; 2002WO-US04284.

XX XX

PR 15-FEB-2001; 2001US-268999P.

XX XX

PA (DIAD-) DIADEXUS INC.

XX XX

PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;

PI PI Sun Y, Liu C;

XX XX

DR WPI; 2002-713345/77.

XX XX

PT New isolated breast specific nucleic acid molecules and polypeptides,

PT useful for identifying, diagnosing, monitoring, staging, imaging and

PT treating breast cancer and non-cancerous disease states in breast

PT tissue

XX XX

PS Claim 1; Page 193-194; 254pp; English.

XX XX

CC The invention relates to human breast specific nucleic acids (I)

CC comprising:

CC (a) a sequence encoding any one of 95 protein sequences

CC (ABP66614-ABP66708);

CC (b) any one of 115 polynucleotide sequences (ABV83558-ABV83672);

CC (c) a molecule that selectively hybridizes to (a) or (b);

CC (d) a molecule having at least 60% sequence identity to (a) or (b).

CC The breast specific nucleic acid molecules, polypeptides and antibodies

CC are useful for identifying, diagnosing, monitoring, staging, imaging and

CC treating breast cancer and non-cancerous disease states in breast tissue.

CC They are also useful for producing transgenic animals and cells and

CC producing engineered breast tissue for treatment and research. The

CC transgenic animals are useful as animal model systems used in elaborating

CC the biological function of the polypeptides, studying conditions and/or

CC disorders associated with aberrant expression and in screening for

CC compounds effective in ameliorating the conditions. The polynucleotides

CC are useful for gene therapy and in vaccines.

XX XX

SQ XX Sequence 430 BP; 135 A; 97 C; 103 G; 95 T; 0 other;

Query Match 8.2%; Score 43.2; DB 24; Length 430;

Best Local Similarity 60.0%; Pred. No. 0.93;

Matches 72; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 50 TTTTGTCTCTCTTATCTATCTTTATTTATTTCCCAATTTTCTTACCAATATTTCTT 109

DB 52 TTTTCTTCTCTCTTATCTATCTTTATTTATTTTCTTCTTCTTCTTCTTCTTCTT 111

QY 110 TGCTAATCAATAAATTATCAAAAGAAAAAATCTGAAAGCAACGCTTCGAAAAAAGGAAA 169
Db 112 TTAAAAAGGGGAATTATAAAAAATTTAAAAAAGGTTTGACCCCAAAAAAATTTAAA 171

Search completed: February 14, 2004, 14:10:12
Job time : 285 secs

ALIGNMENTS

RESULT 1

US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
US-07-867-106-2

Query Match 9.3%; Score 49.2; DB 1; Length 5852;
Best Local Similarity 57.0%; Pred. No. 0.00077;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 12 TGTAGTCGGCTGATTATATGCTGATTATGCGGTGATTGCTTCCTCTTTTACTTT 71
DB 5767 TTTAATATTGTTATGTTATTTTATATATGTTATTTGTTGTTGTTTACTT 5708
QY 72 TTATTATCCCAATTTCTTAAGCAATATTTCTTGCTAATCAATFAAATATCAAA 131
DB 5707 ATATTCTATTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTA 5648
QY 132 AGAAAAAAGCTGAAGCAAGCTGAAAAAGGAAA 169
DB 5647 AAAAAAATAATTTAAATTTAAAAAATAAATAAATAAATAAATAAATAA 5610

RESULT 2

US-09-904-615-66/c
; Sequence 66, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032PI
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-66

Query Match 8.4%; Score 44.4; DB 4; Length 664;
Best Local Similarity 52.8%; Pred. No. 0.006;
Matches 93; Conservative 1; Mismatches 82; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGCGGTGATTGCTTCCTCTTTTACTTTTATTTATCCCAA 86
DB 641 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 582
QY 87 TTTTCTTTAAGCAATATTTCTTTGCTAATCAATAAATATCAAAAGAAAAAACTGA 146
DB 581 ATTTCTGGAAGTTAAAGTAGATACAGCAATATACCAAAAAAAGAAAAAGACA 522
QY 147 AAGCAACCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGAAGTTGTA 202
DB 521 AAAAACTCACATAATAAATTTTACACTATGAAGTACACATTTGGAATTTGAA 466

RESULT 3

US-08-703-947-1/c
; Sequence 1, Application US/08703947
; Patent No. 5788962
; GENERAL INFORMATION:
; APPLICANT: Wise, Kim S.
; APPLICANT: McIntosh, Mark A.
; TITLE OF INVENTION: DNA Sequences Coding for Mycoplasma
; TITLE OF INVENTION: Hyopneumoniae Surface Antigens,
; TITLE OF INVENTION: Corresponding Proteins and Use in
; TITLE OF INVENTION: Vaccines and Diagnostic Procedures
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace J. Fishel
; STREET: 929 Fee Fee Road, Suite 100
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63043
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 KB storage
; COMPUTER: Hewlett-Packard Vectra
; OPERATING SYSTEM: MS-DOS Version 3.3


```
Db      711 TAAATATAAATGTATAA 727
RESULT 8
US-09-904-615-66
; Sequence 66, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-66
Query Match      7.5%; Score 39.6; DB 4; Length 664;
Best Local Similarity 54.2%; Pred. No. 0.12;
Matches 78; Conservative 1; Mismatches 65; Indels 0; Gaps 0;
Qy      26 ATTTATGCTGATTTATGGGATTTGCTTCTTTTATCTTTTATCTTTTATTTATTTCCCAA 85
Db      505 ATTATGTGAGTTTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 564
Qy      86 ATTTTCTTAAGCAAAATATTTCTTTGCTAATCAATAAATATATCAAAAGAAAAAACTG 145
Db      565 CTTTAACTCCAGAAATAAACGTTATATATGGAAGAAAAAAAGAAAAAAAGAAAAAA 624
Qy      146 AAAGCAAGCTTGAAGAAAGGAAA 169
Db      625 AAAAAAGAAAAAGAAAAAGAAAAAGAAA 648
RESULT 9
US-09-475-316A-20/c
; Sequence 20, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Saranen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
; FILE REFERENCE: WSUR-1-13793
; CURRENT APPLICATION NUMBER: US/09/475,316A
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: PCT/US97/20391
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
```

```
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Thuja plicata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(591)
US-09-475-316A-20
Query Match      7.4%; Score 39.4; DB 3; Length 873;
Best Local Similarity 52.8%; Pred. No. 0.16;
Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
Qy      28 TTATATGCTGATTTATGGGATTTTGTCTTCTTTTATCTTTTATTTATTTTCCCAAT 87
Db      873 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 814
Qy      88 TTTTCTTAAGCAAAATATTTCTTCTTAATCAATAAATATCAAAAGAAAAAACTGAA 147
Db      813 TAAACAGCCAAAATACATTTATCTCAACTTACAAGACTCTCCAATATCTTAAACAATTATC 754
Qy      148 AGCAACGCTTGAAAAAGAAAGTTAGCCCTTATCGGGTATA 188
Db      753 AGTAACACTCTAGAGTGTGATGTTAAACCTCAGAGGAAA 713
RESULT 10
US-09-149-476-208/c
; Sequence 208, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
```



```

; EARLIER APPLICATION NUMBER: 60/056, 908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048, 864
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057, 650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056, 884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057, 669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049, 610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061, 060
; EARLIER FILING DATE: 1997-10-02

```

RESULT 11

```

00000001 11
; Sequence 24, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; EARLIER FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-03-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583

```

```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 7.3%; Score 38.8; DB 4; Length 2323;
Best Local Similarity 47.8%; Pred. No. 0.36;
Matches 109; Conservative 1; Mismatches 118; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTTAANGGTTGATTTTGGTCTCTCTTATATCTTTTATCTTTTATCCAAA 86
Db 2291 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTC 2232

QY 87 TTTTCTTTAAGCAATATTTCTTGTCTATCAATAATTTATCAAGAGAAAAAACTGA 146
Db 2231 ACTCTGATAAATTTATTAATAAAGAACAGCATCAGTTGTACATAGAATAATCATTTG 2172

QY 147 AAGCAACGCTCGAAAAAGAAAGTTAGCGCTTATCGGTGTATTTTGGAGTTGTAAAAAT 206
Db 2171 CACTAAGCATGTTACTAGACAGACAGCAGTTTAAATGTTATTTTAAAGTCACAAAAT 2112

QY 207 ACTACRTGTTCTCTTCTAAGTCCCACTCTCTGTCTTTCTTTTGAGCAGG 254
Db 2111 ATATCGGTTAAGGCACACTATAATCTCTGTCATGTTTTCATTAATAATG 2064

RESULT 12
US-09-212-971-3
; Sequence 3, Application US/09212971B
; Patent No. 6107041
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Teang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009002
; CURRENT APPLICATION NUMBER: US/09/212,971B
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 60/017,354
; EARLIER FILING DATE: 1996-04-26
; EARLIER APPLICATION NUMBER: 60/030,590
; EARLIER FILING DATE: 1996-11-14
; EARLIER APPLICATION NUMBER: 08/800,929
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (4623)...(4623)

```

```
; OTHER INFORMATION: n can be any nucleotide
; FEATURE:
; NAME/KEY: variation
; LOCATION: (4622)...(4622)
; OTHER INFORMATION: n can be any nucleotide
US-09-212-971-3

Query Match          7.3%; Score 38.8; DB 3; Length 5232;
Best Local Similarity 62.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 34 GCTGATTTATGGGTGATTTTCTTCTTCTTATATCTTTTATTTATTTCCCAAAATTTTCT 93
Db 2334 GGTTTATAGGGGCTTTTCTACTTTCTACTTTTCTACTTTTGTCTGTCGAATTTTTA 2393

Qy 94 TAAGCAAAATATTTCTTCTGCTAATCAATAAATATCAAA 131
Db 2394 TAAGTATGTATTACTTTTGTAAATCAGAATTTTAGAAA 2431

RESULT 13
US-08-800-929A-3
; Sequence 3, Application US/08800929A
; Patent No. 6133437
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF
; TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,929A
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,590
; FILING DATE: 14-NOV-1996
; APPLICATION NUMBER: 60/017,354
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 07891/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5232 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Other
```

```
; LOCATION: 1...5232
; OTHER INFORMATION: N at 4622 and 4633 can be A, G, T or C.
US-08-800-929A-3

Query Match          7.3%; Score 38.8; DB 3; Length 5232;
Best Local Similarity 62.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 34 GCTGATTTATGGGTGATTTTCTTCTTCTTATATCTTTTATTTATTTCCCAAAATTTTCT 93
Db 2334 GGTTTATAGGGGCTTTTCTACTTTCTACTTTTCTACTTTTGTCTGTCGAATTTTTA 2393

Qy 94 TAAGCAAAATATTTCTTCTGCTAATCAATAAATATCAAA 131
Db 2394 TAAGTATGTATTACTTTTGTAAATCAGAATTTTAGAAA 2431

RESULT 14
US-09-617-053A-3
; Sequence 3, Application US/09617053A
; Patent No. 6300492
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; PRIOR FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (4623)...(4623)
; OTHER INFORMATION: n can be any nucleotide
; NAME/KEY: variation
; LOCATION: (4622)...(4622)
; OTHER INFORMATION: n can be any nucleotide
US-09-617-053A-3

Query Match          7.3%; Score 38.8; DB 4; Length 5232;
Best Local Similarity 62.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 34 GCTGATTTATGGGTGATTTTCTTCTTCTTATATCTTTTATTTATTTCCCAAAATTTTCT 93
Db 2334 GGTTTATAGGGGCTTTTCTACTTTCTACTTTTCTACTTTTGTCTGTCGAATTTTTA 2393

Qy 94 TAAGCAAAATATTTCTTCTGCTAATCAATAAATATCAAA 131
Db 2394 TAAGTATGTATTACTTTTGTAAATCAGAATTTTAGAAA 2431

RESULT 15
US-08-545-196B-10/c
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
```



```
Db 15117 GGTGTTGTTTTTGGCTCTTTTTTTTTTTTTTCCAGTATGGAATCCATCTGTTCACGACC 15058
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
Qy 85 AATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAATAAACT 144
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
Db 15057 ATTGTTGAAAGACTTCTTTCTATTCATATTTATTTGGGAGTGAAAAAAAAC 14998
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
Qy 145 GAAACAAAGCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGGAAAGTTGTAAA 204
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
Db 14997 AAAAGAAAACTTTGCAATAACAGGTTTCACATAAAACCAATTTCTTGCTGTGGAGG 14938
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
Qy 205 AT 206
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
Db 14937 AT 14936
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11

RESULT 18
US-09-410-464-5
; Sequence 5, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
; APPLICANT: Straus et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410,464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1999-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 5656
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-5
Query Match 7.1%; Score 37.8; DB 4; Length 5656;
Best Local Similarity 58.4%; Pred. No. 1;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
Qy 60 TCTTTATATCTTTTATTTATTTCCCAATTTTCTTTAAGCAATATTTCTTTGCTATCAAA 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4024 TTATTTAGAAATATATTAATAATATTTTATATTTTAAATTTATTTTAAATTTAA 4083
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 120 TAAATTTCAAAGAAAAAACTGAAGCAACGCTTGAAAAAGGAAAGTT 172
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4084 TATATTAAATAATATAAAATACTGAAAAATAATTTTAAAAATAATTTT 4136
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 19
US-09-149-476-66
; Sequence 66, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```



```

Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 42 ATGGGTGATTTCCTTCCTTTTATATCTCTTTTATTTATTCCTCAAAATTTCTTTAAGCAA 101
Db 3671 ATTTGTTGTTTTCTTTTACATGACTTTTATAGTTATTCGGTAGTTATTGTATAATACT 3730
Qy 102 TATTCTTTCTTAATCAATAAATATCAAAAGAAAAAATCTGAAGCAACGGCTTGAA 161
Db 3731 GATAACGATCATATATACACACTTTGTTCACAAAAAATTTTTTTTTTTTTTTTTT 3790
Qy 162 AAAG 165
Db 3791 AAAG 3794

RESULT 22
US-09-601-198-165/c
; Sequence 165, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Casseil, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 4344
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-165

```

```

      88  TTTTCTTAAAGCAATAATTTCTTTTGCTAATCAATAAATTTATCAAAAGAAAAAAAACCTGAA 147
      3218  AAATCATTAATACTACTTCATCATCAATAATCTTGATTAATTTTTTTCATCATCTTATAATAAATTTA 3159
      148  AGCAACGCTTGAAAAAAGGA 167
      3158  TATGATTCCTTTTAAATAATGA 3139

RESULT 23
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, M
; Patent No. 6503729
; FILE REFERENCE: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3

```

```
;/ SOFTWARE: PatentIn version 3.1
;/ SEQ ID NO 1
;/ LENGTH: 1664976
;/ TYPE: DNA
;/ ORGANISM: Methanococcus jannaschii
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (28222)..(28222)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (28257)..(28258)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (84773)..(84773)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (84808)..(84808)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (84812)..(84812)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (98120)..(98120)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (98159)..(98159)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (98239)..(98239)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (98266)..(98266)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (98343)..(98343)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (103998)..(103998)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (148948)..(148948)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (163385)..(163385)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (191989)..(191989)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (191995)..(191995)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (231980)..(231980)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234187)..(234187)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234220)..(234220)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234814)..(234814)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (309398)..(309398)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (309418)..(309418)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (312837)..(312837)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (312993)..(312993)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (319226)..(319226)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (559167)..(559167)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (559241)..(559241)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (600992)..(600992)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (622708)..(622708)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (657081)..(657081)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (657203)..(657203)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (674435)..(674435)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (682442)..(682442)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (713652)..(713652)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (741684)..(741684)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (779455)..(779455)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (779676)..(779676)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (855539)..(855539)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (871619)..(871619)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1084830)..(1084830)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1096846)..(1096846)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1119881)..(1119881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1130881)..(1130881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1310988)..(1310988)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1313224)..(1313224)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1349473)..(1349473)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1349491)..(1349491)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1470091)..(1470091)
```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match
Best Local Similarity 7.1%; Score 37.6; DB 4; Length 1664976;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 24 TGATTATATCTGATTTTGGTGATTTTGGTCTCTTTATACTTTATTTATTTCC 83
Db 1332304 TAAATCAACATCTACTCATTTAGTTTATATAATTCCTTGCNAACATTTTAAACATCC 1332363
Qy 84 AAATTTTCTTAAGCAATATTTCTTGTCTATCAATAAATATCAAAAGAAAAAACA 143
Db 1332364 ACTGTGTTTTTAGATTAATTTCTTTGTTAATAATAAGCTTACACATAAACAAT 1332423
Qy 144 TGAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGGAAAGTTGPA 203
Db 1332424 TGAGTAGTATATCTTGAAGCAGCAAAATCATAAATTCGCTATTATAAGATTTCTGA 1332483

RESULT 24
US-09-369-247-23/c
Sequence 23, Application US/09369247
Patent No. 6569992
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 44 Human Secreted Proteins
FILE REFERENCE: P2024P1
CURRENT APPLICATION NUMBER: US/09/369,247
EARLIER FILING DATE: 1999-08-05
EARLIER APPLICATION NUMBER: 60/074,118
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,157
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,137
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,341
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,141
EARLIER FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 1492
TYPE: DNA
ORGANISM: Homo sapiens
US-09-369-247-23

Query Match
Best Local Similarity 7.0%; Score 37.2; DB 4; Length 1492;
Matches 57; Conservative 2; Mismatches 35; Indels 0; Gaps 0;
Qy 39 TTTATGGGTGATTTTGGCTCTCTTTATACTTTATTTATTTATTTCCCAATTTTCTTAAGC 98
Db 1484 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1425
Qy 99 AAATATTTCTTGTCTAATCAATAAATATCAAAA 132

Db 1424 TTTTATTTTGTGAAAAACCAATAATTTATCAAAA 1391

RESULT 25

US-07-867-106-2
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Peeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match
Best Local Similarity 7.0%; Score 37.2; DB 1; Length 5852;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
Qy 27 TTTATATGCTGATTTATGGGTGATTTTGGCTCTCTTTATCTTTATCTTTATTTATTTCCCAAA 86
Db 5594 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5653
Qy 87 TTTTCTTACCAATATTTCTTTGCTAATCAATAATATCAAAAGAAAAAACTGA 146
Db 5654 TTTTATTTTATTTAAATTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTT 5713
Qy 147 AA 148
Db 5714 AA 5715

RESULT 26

US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-545-196B-10

Query Match 7.0%; Score 37; DB 3; Length 1582;
Best Local Similarity 56.0%; Pred. No. 0.95;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 45 GGTGATTTGGCTTCCTCTTATCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATAT 104
Db 1420 GTTCATTGCTACTGTTTTTCTCTCTATATGTTTAAAGTATATAATAAAATATTT 1479
QY 105 TTCTTTGCTAAATCAATAATTTATCAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 164
Db 1480 AATTTTTTTTAAAAAATAATTTATCAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 1539
QY 165 GGAAA 169
Db 1540 AAAAA 1544

RESULT 27

US-08-545-196B-12
; Sequence 12, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH

STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,196B
FILING DATE: 19-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FARACI, C. J.
REGISTRATION NUMBER: 32,350
REFERENCE/DOCKET NUMBER: 2121-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-545-196B-12

Query Match 7.0%; Score 37; DB 3; Length 1582;
Best Local Similarity 56.0%; Pred. No. 0.95;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 45 GGTGATTTGGCTTCCTCTTATCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATAT 104
Db 1420 GTTCATTGCTACTGTTTTTCTCTCTATATGTTTAAAGTATATAATAAAATATTT 1479
QY 105 TTCTTTGCTAAATCAATAATTTATCAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 164
Db 1480 AATTTTTTTTAAAAAATAATTTATCAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 1539
QY 165 GGAAA 169
Db 1540 AAAAA 1544

RESULT 28

US-08-747-221B-36
; Sequence 36, Application US/08747221B
; Patent No. 6063610
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,221B
; FILING DATE: No. 6063610ember 12, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1594
US-08-747-221B-36

Query Match 7.0%; Score 37; DB 3; Length 2007;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 24 TGATTATATGCTGATTATGGGTGATTTCCTTCCTTCTTATATCTTTTATTTATTTCC 83
Db 1883 TTATTTATTTGATATATTTACCATCTTTGTATCATATTTGCTTTTATTTTTCATTTT 1942
Qy 84 AAATTTTCTTAAGCAATATTTCTTCCTATCAATAAATATCAAAAGAAAAAAC 143
Db 1943 TTTTATTTCAATATATTTGTTTTTATAAAAAATTTTATAAAAAATTTTATTTT 2002
Qy 144 TGA 148
Db 2003 AAAAA 2007

RESULT 29
US-08-747-221B-38/c
; Sequence 38, Application US/08747221B
; Patent No. 6063610
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,221B
; FILING DATE: No. 6063610ember 12, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 nucleotides
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-747-221B-38
Query Match 7.0%; Score 37; DB 3; Length 2007;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 24 TGATTATATGCTGATTATGGGTGATTTCCTTCCTTCTTATATCTTTTATTTATTTCC 83
Db 125 TTATTTATTTGATATATTTACCATCTTTGTATCATATTTGCTTTTATTTTTCATTTT 66
Qy 84 AAATTTTCTTAAGCAATATTTCTTCCTATCAATAAATATCAAAAGAAAAAAC 143
Db 65 TTTTATTTCAATATATTTGTTTTTATAAAAAATTTTATAAAAAATTTTATTTT 6
Qy 144 TGA 148
Db 5 AAAAA 1

RESULT 30
US-09-005-051-36
; Sequence 36, Application US/09005051
; Patent No. 6291222
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,051
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/747,221
; FILING DATE: No. 6291222ember 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1594
US-09-005-051-36
Query Match 7.0%; Score 37; DB 3; Length 2007;
Best Local Similarity 56.0%; Pred. No. 1.1;

	Query Match	7.0%	Score 37	DB 3	Length 2007	
	Best Local Similarity	56.0%	Pred. No. 1.1	Mismatches 55	Indels 0	Gaps 0
	Matches	70	Conservative			
QY	24	TGATTTATATGCTGATTTATGCGGTGATTTTTCCTCTCTTTACCTTTATTTATTTCC	83			
DB	125	TTATTTATATTTGCATATTTATGACATCTTTGTATCATATTTGCTTTATTTTTCATTTT	66			
QY	84	AAATTTTTCTTAGCGAAATATTTCTTTTGCTAAATCAATAAATTTACAAAGAAAAA	143			
DB	65	TTTTTTATTTCAATATATTTGTTTTTTTATAAAAAA	6			

```

Query Match      7.0%; Score 37; DB 1; Length 2334;
Best Local Similarity 53.9%; Pred. No. 1.1;
Matches 76; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY      28  TTATATGCTGATTTATGGGTGATTTTGCCTTCCTTTATACCTTTTATTTATTCCTCAAT 87
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      458  TTATATCAATGTTTATATAAATATTCCTTCCTTATTTTACTTTTTTTTTTTTTTTT 517
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      88  TTTTCTTAAGCAAAATATTTCTTTGCTAAATCAATAAATATCAAAAGAAAAAAAACCTGAA 147
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      518  TTTTCTTTGTTAAATATTTATGTCATAGTAAATTTAAGTTAAACCACCAACCTAACTTAT 577
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      148  AGCAACGCTTGAAAAAGGAA 168
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      578  AAAACCGAAAGGAATTAGTAA 598
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 33
US-09-234-245-1
; Sequence 1, Application US/09234245
; Patent No. 6509457
; GENERAL INFORMATION:
; APPLICANT: Thomas Jefferson University
; APPLICANT: Raphael Rubin
; APPLICANT: Manorama Tewari

```


Qy	100	ATAATTCTCTTGGTAATCAATAAATATCAAAAGAAAAAANAACATGAAGCAACGCTTGA	159
Db	14571	CANAATTCATTTAGAAATTTTGAAGAGATTGTATAGGTTTAAACCTCTCAATTTCAATTACA	14630
Qy	160	AAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAAATACTACRTGTTCTC	219
Db	14631	GAGTGGAAACCCAGTCTTATATACAAATCTTTTGATTTTTTTTTTACAGGAGTTTTTCAA	14690
Qy	220	TTCT	223
Db	14691	TTGT	14694

```

RESULT 36
US-09-489-847-112/c
; Sequence 112, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0

```

RESULT 38
US-09-328-352-642
; Sequence 642, Application US/09328352

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match	6.8%	Score 36.2;	DB 4;	Length 318;
Best Local Similarity	57.5%;	Pred. NO. 0.77;		
Matches 65;	Conservative	0;	Mismatches 48;	Indels 0;
				Gaps 0;

Qy 147 AAGCAACGCTTGAAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGAAGTT 199

nb 121 GATCAGCTTATGAAAAAACTAATGGTGAGCATTTTCGCTTCTTTTTTTTTTCACTTT 173

RESULT 39
US-09-453-323-1
; Sequence 1, Application US/09453323
; Patent No. 6441272

```

/ GENERAL INFORMATION:
/ APPLICANT: Ye, Zheng-Hua
/ TITLE OF INVENTION: MODIFICATION OF LIGNIN CONTENT AND COMPOSITION IN
/ TITLE OF INVENTION: PLANTS
/ TITLE OF INVENTION: PLANTS
/ FILE REFERENCE: 235,00090101

```

Qy	100	ATAATTTCTCTTGGTAATCAATAAATATCAAAAGAAAAAANAACATGAAGCAACGCTTGA	159
Db	14571	CANAATTCATTTAGAAATTTTGAAGAGATTGTATAGGTTTAAACCTCTCAATTTCAATTACA	14630
Qy	160	AAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAAATACTACRTGTCTC	219
Db	14631	GAGTGGAAACCCAGTCTTATATACAAATCTTTTGATTTTTTTTTTACAGGAGTTTTTCAA	14690
Qy	220	TTCT	223
Db	14691	TTGT	14694

```

RESULT 36
US-09-489-847-112/c
; Sequence 112, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0

```

RESULT 37
US-08-894-731-2
; Sequence 2, Application US/08894731
; Patent No. 6084089
; GENERAL INFORMATION:


```
; CURRENT APPLICATION NUMBER: US/09/453,323
; CURRENT FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/110,676
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-09-453-323-1

Query Match
Best Local Similarity 6.8%; Score 36; DB 4; Length 1507;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 46 GTGATTTTGGCTTCCTTTTACATTTTATTTCCCAAAATTTTCTTAAGCAATATT 105
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1357 GTGATTTGAGTTTGAGCTGTATTCGGTGTATTCCTCAATTCCTCCCTAAGCAAGATAT 1416
QY 106 TCTTTGCTAATCAATAATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAG 165
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1417 TAGCAGATGATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1476

QY 166 GAAA 169
DB |||||
1477 AAAA 1480

RESULT 40
US-08-821-994-59
; Sequence 59, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier RP
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-59

Query Match
Best Local Similarity 52.7%; Score 36; DB 3; Length 1577;
Matches 78; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 22 GCTGATTTATATGCTGATTTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTC 81
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1426 GATGATTTACTTTATAGCTGTTGTTGGATATGATATATTAGTCTCTTTATTTGGATGATA 1485
QY 82 CCAAAATTTTCTTAAGCAAAATTTTCTTTGCTTAATCAATAAATTAATCAAAAGAAAAA 141
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1486 CAACTTTTGAATCAATAAAGTTTACTTGCAGGACACATAAAAAAATAAAAAAATA 1545
QY 142 ACTGAAGCAACGCTTGAAAAAGAAA 169
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1546 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1573

RESULT 41
US-09-293-322C-8
; Sequence 8, Application US/09293322C
```

```
; Patent No. 6232110
; GENERAL INFORMATION:
; APPLICANT: Pallas, David C
; APPLICANT: Du, Xianxing
; TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,
; Patent No. 6232110
; TITLE OF INVENTION: Recombinant DNA Molecules and Methods
; FILE REFERENCE: 105-97
; CURRENT APPLICATION NUMBER: US/09/293,322C
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/082,202
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2409)
; OTHER INFORMATION: N is A, T, G or C.
US-09-293-322C-8

Query Match
Best Local Similarity 6.8%; Score 36; DB 3; Length 2409;
Matches 60; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 70 TTTTATTATTCCTCCAAATTTTCTTAAGCAAAATTTTCTTTGCTTAATCAATAAATTATCA 129
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2300 TCTTTGTTATTATGATCTTGTTTAAAGAAAAAATAAATATCTCCCAACCTTTTAAAAA 2359

QY 130 AAAGAAAAAATACTGAAAGCAACGCTTGAAAAAGGAAA 169
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2360 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2399

RESULT 42
US-09-839-497A-8
; Sequence 8, Application US/09839497A
; Patent No. 6528295
; GENERAL INFORMATION:
; APPLICANT: Pallas, David C.
; APPLICANT: Du, Xianxing
; TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,
; Patent No. 6528295
; TITLE OF INVENTION: Recombinant DNA Molecules and Methods
; FILE REFERENCE: Docket No. 6528295 105-97A
; CURRENT APPLICATION NUMBER: US/09/839,497A
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/082,202
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/293,322
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2409)
; OTHER INFORMATION: N is A, T, G or C.
US-09-839-497A-8

Query Match
Best Local Similarity 6.8%; Score 36; DB 4; Length 2409;
Matches 60; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 70 TTTTATTATTCCTCCAAATTTTCTTAAGCAAAATTTTCTTTGCTTAATCAATAAATTATCA 129
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2300 TCTTTGTTATTATGATCTTGTTTAAAGAAAAAATAAATATCTCCCAACCTTTTAAAAA 2359
```


EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 76
LENGTH: 519
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (13)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (24)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE

LOCATION: (35)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (44)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-76
Query Match
Best Local Similarity 59.0%; Score 35.6; DB 4; Length 519;
Matches 59; Conservative 1; Mismatches 40; Indels 0; Gaps 0;
Oy 46 GTGATTTTGGCTTCCTTTTATATCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAATATT 105
Db 497 GAGTTTTTTTTTWTYTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAACATTTATT 438
Oy 106 TCITTTGCTAATCAATTAATATCAAAAGAAAAAACTG 145
Db 437 TATCTACTGTACAAATATTTTACATCATCATGCTGCAACTG 398
Search completed: February 14, 2004, 15:19:13
Job time : 81 secs

THIS PAGE BLANK (USPTO)